

ON THE DEVELOPMENT OF SPACE-NUMBER RELATIONS: LINGUISTIC AND COGNITIVE DETERMINANTS, INFLUENCES, AND ASSOCIATIONS

EDITED BY: Hans-Christoph Nuerk, Krzysztof Cipora, Frank Domahs and
Maciej Haman

PUBLISHED IN: Frontiers in Psychology





frontiers

Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence.

The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714

ISBN 978-2-88963-588-7

DOI 10.3389/978-2-88963-588-7

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

ON THE DEVELOPMENT OF SPACE-NUMBER RELATIONS: LINGUISTIC AND COGNITIVE DETERMINANTS, INFLUENCES, AND ASSOCIATIONS

Topic Editors:

Hans-Christoph Nuerk, University of Tübingen, Germany

Krzysztof Cipora, University of Tübingen, Germany

Frank Domahs, University of Erfurt, Germany

Maciej Haman, University of Warsaw, Poland

Citation: Nuerk, H.-C., Cipora, K., Domahs, F., Haman, M., eds. (2020). On the Development of Space-Number Relations: Linguistic and Cognitive Determinants, Influences, and Associations. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88963-588-7

Table of Contents

05	<i>Editorial: On the Development of Space-Number Relations: Linguistic and Cognitive Determinants, Influences, and Associations</i> Krzysztof Cipora, Maciej Haman, Frank Domahs and Hans-Christoph Nuerk
10	<i>Developmental Dyscalculia and Automatic Magnitudes Processing: Investigating Interference Effects Between Area and Perimeter</i> Hili Eidlin-Levy and Orly Rubinsten
24	<i>Are Books Like Number Lines? Children Spontaneously Encode Spatial-Numeric Relationships in a Novel Spatial Estimation Task</i> Clarissa A. Thompson, Bradley J. Morris and Pooja G. Sidney
34	<i>Why Numbers are Embodied Concepts</i> Martin H. Fischer
37	<i>Commentary: The Mental Representation of Integers: An Abstract-to-concrete Shift in the Understanding of Mathematical Concepts</i> Melinda A. Mende, Samuel Shaki and Martin H. Fischer
41	<i>Developmental Changes in the Effect of Active Left and Right Head Rotation on Random Number Generation</i> Charlotte Sosson, Carrie Georges, Mathieu Guillaume, Anne-Marie Schuller and Christine Schiltz
55	<i>From Innate Spatial Biases to Enculturated Spatial Cognition: The Case of Spatial Associations in Number and Other Sequences</i> Koleen McCrink and Maria Dolores de Hevia
59	<i>Magnitude or Multitude – What Counts?</i> Martin Lachmair, Susana Ruiz Fernández, Korbinian Moeller, Hans-Christoph Nuerk and Barbara Kaup
66	<i>Numerical Affordance Influences Action Execution: A Kinematic Study of Finger Movement</i> Rosa Rugani, Sonia Betti and Luisa Sartori
76	<i>How Deep is Your SNARC? Interactions Between Numerical Magnitude, Response Hands, and Reachability in Peripersonal Space</i> Johannes Lohmann, Philipp A. Schroeder, Hans-Christoph Nuerk, Christian Plewnia and Martin V. Butz
91	<i>Spatial-Numerical Associations Enhance the Short-Term Memorization of Digit Locations</i> Catherine Thevenot, Jasinta Dewi, Pamela B. Lavenex and Jeanne Bagnoud
97	<i>Spatial and Verbal Routes to Number Comparison in Young Children</i> Francesco Sella, Daniela Lucangeli and Marco Zorzi
106	<i>Implicit and Explicit Number-Space Associations Differentially Relate to Interference Control in Young Adults With ADHD</i> Carrie Georges, Danielle Hoffmann and Christine Schiltz
122	<i>The Connection Between Spatial and Mathematical Ability Across Development</i> Christopher J. Young, Susan C. Levine and Kelly S. Mix

- 129 ***A Taxonomy Proposal for Types of Interactions of Language and Place-Value Processing in Multi-Digit Numbers***
Julia Bahnmüller, Hans-Christoph Nuerk and Korbinian Moeller
- 136 ***A Mental Odd-Even Continuum Account: Some Numbers May be “More Odd” Than Others and Some Numbers May be “More Even” Than Others***
Lia Heubner, Krzysztof Cipora, Mojtaba Soltanlou, Marie-Lene Schlenker, Katarzyna Lipowska, Silke M. Göbel, Frank Domahs, Maciej Haman and Hans-Christoph Nuerk
- 152 ***Corrigendum: A Mental Odd-Even Continuum Account: Some Numbers May be “More Odd” Than Others and Some Numbers May be “More Even” Than Others***
Lia Heubner, Krzysztof Cipora, Mojtaba Soltanlou, Marie-Lene Schlenker, Katarzyna Lipowska, Silke M. Göbel, Frank Domahs, Maciej Haman and Hans-Christoph Nuerk
- 153 ***The Developing Mental Number Line: Does its Directionality Relate to 5- to 7-Year-Old Children’s Mathematical Abilities?***
Lauren S. Aulet and Stella F. Lourenco
- 168 ***The Developmental Trajectory of the Operational Momentum Effect***
Pedro Pinheiro-Chagas, Daniele Didino, Vitor G. Haase, Guilherme Wood and André Knops
- 182 ***Compatibility Between Physical Stimulus Size – Spatial Position and False Recognitions***
Seda Dural, Birce B. Burhanoglu, Nilsu Ekinci, Emre Gürbüz, İdil U. Akın, Seda Can and Hakan Çetinkaya
- 194 ***The Use of Local and Global Ordering Strategies in Number Line Estimation in Early Childhood***
Jaccoline E. Van 't Noordende, M. J. M. Volman, Paul P. M. Leseman, Korbinian Moeller, Tanja Dackermann and Evelyn H. Kroesbergen
- 210 ***Differential Development of Children’s Understanding of the Cardinality of Small Numbers and Zero***
Silvia Pixner, Verena Dresen and Korbinian Moeller
- 221 ***Variability in the Alignment of Number and Space Across Languages and Tasks***
Andrea Bender, Annelie Rothe-Wulf and Sieghard Beller
- 240 ***Development of a Possible General Magnitude System for Number and Space***
Karin Kucian, Ursina McCaskey, Michael von Aster and Ruth O’Gorman Tuura
- 253 ***Commentary : The Developmental Trajectory of the Operational Momentum Effect***
Martin H. Fischer, Alex Miklashevsky and Samuel Shaki
- 256 ***Linear Spatial–Numeric Associations Aid Memory for Single Numbers***
John Opfer, Dan Kim, Christopher J. Young and Francesca Marciari
- 268 ***English and Chinese Children’s Performance on Numerical Tasks***
Ann Dowker and Anthony M. Li
- 279 ***Response: Commentary: The Developmental Trajectory of the Operational Momentum Effect***
Daniele Didino, Pedro Pinheiro-Chagas, Guilherme Wood and André Knops



Editorial: On the Development of Space-Number Relations: Linguistic and Cognitive Determinants, Influences, and Associations

Krzysztof Cipora^{1,2}, Maciej Haman³, Frank Domahs⁴ and Hans-Christoph Nuerk^{1,2*}

¹ Department of Psychology, University of Tübingen, Tübingen, Germany, ² LEAD Research Network, University of Tübingen, Tübingen, Germany, ³ Faculty of Psychology, University of Warsaw, Warsaw, Poland, ⁴ Department of Linguistics, University of Erfurt, Erfurt, Germany

Keywords: Spatial-Numerical Association, number processing, spatial biases, cognitive development, Spatial-Numerical Association taxonomy

Editorial on the Research Topic

On the Development of Space-Number Relations: Linguistic and Cognitive Determinants, Influences, and Associations

OPENING REMARKS

Tight, bidirectional links between mathematical cognition and spatial processing are documented (Cipora et al., 2015). For instance, there is a relationship between mathematical and spatial development (Young et al.). Individuals who perform well on spatial tasks perform well on math tasks as well (Mix et al., 2016, for a review). Individuals with impairments in mathematical processing are more prone to interference on spatial tasks, even when such tasks do not contain numerical components (Eidlin-Levy and Rubinsten). Apart from correlations between spatial and numerical skills and deficits, there is a broad range of phenomena linking numerical and spatial processing referred to as “Spatial-Numerical Associations” (SNAs). SNAs are not just correlates of numerical processing and math skills, it is also supposed that they may be the key to hidden, deep properties of numerical representations and processes operating on them.

In this context, the question not only of how SNAs correlate with numerical development but also what their role in this development is arises. Lifetime development of SNAs and their functional role is therefore a focus of this Research Topic. As one can see from this collection of 27 papers, there is a considerable variety of SNAs. Our guidance through this variety is based on a taxonomy of SNAs we have proposed and extended (Patro et al., 2014; Cipora et al., 2015, 2018b).

The major distinction in this taxonomy of SNAs is between *Extension* and *Direction* SNAs. In extension SNAs, numbers are associated with (one- or multi-dimensional) extensions in space: larger numbers are associated with larger extensions. There are two subcategories within extension SNAs. In *Approximate Extension* SNAs “more” in one domain corresponds to “more” in the other one (e.g., in the numerical Stroop task larger font size is associated with larger numerical magnitude, but there is no relationship between specific font size and specific numerical magnitude). In *Exact Extension* SNAs, the accuracy of a relationship between number magnitude and spatial extension is examined. The variable of interest is usually the deviation from the exact isomorphy of numerical and spatial magnitude. For instance, in the number line estimation task the deviation of a marked spatial position on the number line and the position corresponding to the exact numerical magnitude given is analyzed.

OPEN ACCESS

Edited and reviewed by:

Carmelo Mario Vicario,
University of Messina, Italy

*Correspondence:

Hans-Christoph Nuerk
hc.nuerk@uni-tuebingen.de

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 22 January 2020

Accepted: 27 January 2020

Published: 14 February 2020

Citation:

Cipora K, Haman M, Domahs F and
Nuerk H-C (2020) Editorial: On the
Development of Space-Number
Relations: Linguistic and Cognitive
Determinants, Influences, and
Associations. *Front. Psychol.* 11:182.
doi: 10.3389/fpsyg.2020.00182

In case of *Direction SNAs*, relationships between numbers and specific directions in space (left-right, up-down, near-far) are investigated. The association can either be explicit (as revealed in overt controllable actions such as ordering objects in a certain dimension), or implicit (e.g., a reaction time pattern; Dehaene et al., 1993; Nuerk et al., 2005b). Importantly, several aspects of a number can be associated with space: cardinality, ordinality, functions, and place-value structures.

The rationale for such a taxonomy and distinction is conceptual and is not only aimed at emphasizing peculiarities of different tasks used to measure SNAs. SNA types differ in several fundamental aspects: (1) their relationship with math skill, (2) their potential for being used in interventions, and (3) the extent to which they are prone to situated influences (Cipora et al., 2018c). We are glad to see that papers published in the Research Topic cover all SNA types from the taxonomy (except two types, which we theoretically postulated, but could not find any existing study supporting their existence).

DIRECTION IMPLICIT SNAs

Cardinality

This SNA category is the most investigated in the literature and also very well-covered in the current Research Topic. Among other phenomena it considers the SNARC effect (Spatial-Numerical Association of Response Codes; Dehaene et al., 1993; Wood et al., 2008), the hallmark effect showing left-to-right mapping of numerical magnitude representations in Western cultures. Papers published in this volume focusing on this SNA category were either aimed at investigating the foundations of this SNA type in general, or searched for a functional role at different stages of development. In the first group of papers, McCrink and de Hevia outline a new theory on origins of directional SNAs. Their work attempts to integrate opposing views on directionality in SNAs: the nativist (e.g., Rugani et al., 2015; Di Giorgio et al., 2019) and culture related (Shaki et al., 2009; Patro et al., 2016a,b; Patro and Nuerk, 2017). According to their new proposal, left-to-right SNAs are inborn, but in humans they weaken at toddlerhood and then are further (re)shaped by cultural factors. Sosson et al. provide empirical evidence for the developmental trajectory of directional spatial biases in an infrequently used task: random number generation. Adult leftward head movements are related to generating smaller random numbers compared to numbers generated during rightward movements, however, such an effect was not observed in children. SNAs are not only influenced by development *per se*, but also by developmental disorders. Georges et al. looked at the SNARC effect in individuals with ADHD. They found that weaker inhibition capacities were related to a stronger SNARC effect as measured with magnitude-irrelevant parity judgment, and a weaker SNARC effect as measured with a magnitude-relevant magnitude classification task.

Whether a relation between directional SNAs and arithmetic skill exists, is controversial (Cipora et al., 2015, 2018b,c, 2019). Aulet and Lourenco report null, or in one task even negative, correlations between SNA strength and arithmetic skills, corroborating earlier findings that directional SNAs are

at least not consistently linked to math skills. However, this does not imply that directional SNAs cannot be used to enhance mathematical understanding as a symbolic tool. Indeed, Thevenot et al. showed that SNAs can enhance memorization of digits in a memory game. Such findings have potential future educational applications.

Another group of papers were not focused on development or associations with math skills, but on the foundations and measurement of SNAs. In a theoretical contribution, Mende et al. discuss the influence of motoric responses. They suggest that the spatial associations of negative numbers might be not a direct measure of their representation, but rather influenced by response paradigm, namely, if an individual has two horizontally aligned response keys or just one. The influence of situated and embodied factors on SNAs seems to be important (Cipora et al., 2018a). Using a 3D virtual setup, Lohmann et al. demonstrate that the strength of the SNARC effect is modulated by perceived physical proximity between numbers and (virtual) responding hands. Finally, language attributes can also influence directional SNAs. Lachmair et al. contrast the role of magnitude and multitude (singular or plural grammatical number) on vertical SNAs, showing that magnitude is a more robust factor.

In sum, papers in this Research Topic (together with other literature) suggest that directional SNAs are multi-faceted in their development, that their relation to arithmetic skills remains ambiguous and that other domain-general factors like motoric or linguistic factors are relevant in their investigation.

Ordinality

Numerous studies have demonstrated this SNA type, including SNARC-like effects for non-numerical sequences such as days of the week (Gevers et al., 2003) or object position in working memory (van Dijck and Fias, 2011). Dural et al. used a slightly modified paradigm and show a SNARC-like effect for object size in a new memory retrieval paradigm, thereby extending generality of this SNA type across different paradigms.

Functions

In past years we witnessed a vivid discussion about spatial biases in mental arithmetic: One major finding of debate was the Operational Momentum (OM) Effect (Pinhas and Fischer, 2008; Knops et al., 2009). Pinheiro-Chagas et al. investigated its developmental trajectory in a group of children ages 8–12. They observed a monotonic increase in the size of their OM with age and attributed this finding to increasing reliance on the Mental Number Line representation while performing mental calculations and involvement of attentional processes. Fischer et al. comment on this paper emphasizes multifaceted origins of the OM Effect, which cannot only be accounted for solely by attentional processes. Instead, they argue that their AHAB (arithmetic heuristics and biases) model better accounts for OM development (including the reversed OM in 6-year-olds). The authors of the original paper (Didino et al.) refute the criticism of Fischer et al. by (1) providing arguments that the alternative mechanism of logarithmic compression of the mental number line, on which additions and subtractions are performed was not a strawman argument; (2) stating that

heuristics and the attentional account are hardly distinguishable empirically; (3) providing alternative scenarios for which kind of vertical movement should be related to addition and subtraction. At the same time, they call for more precise definitions of the OM and estimation biases induced by the operation sign.

Place Value

In their taxonomy of interactions between languages and place-value processing, Bahnmueller et al. (esp. Table 1) describe a variety of implicit and explicit place-value effects at three place-value levels: place identification, place-value activation, and place-value computation. Place-value processing can be influenced at all of these levels by various linguistic attributes.

DIRECTION EXPLICIT SNAs

Cardinality and Functions

In our taxonomy we postulate the existence of these subcategories, but so far we have not identified any studies investigating them.

Ordinality

Explicit directional SNAs have been mostly investigated in terms of cultural differences in counting. Based on a large-scale cross-cultural study on five different cultures Bender et al. suggest that the mere focus on main effects of cultural attributes falls short, because characteristics of task and paradigm as well as individual differences need to be considered to obtain a more comprehensive picture.

Place-Value

In our taxonomy, we have defined place-value processing as a directional SNA because the correct processing of the spatial direction of digits is necessary to assess place-value magnitude (29 vs. 92). Linguistic factors are known to influence multi-digit number processing (e.g., Nuerk et al., 2005a; Moeller et al., 2015; Bahnmueller et al.). Dowker and Li compare English and Hong Kong (L1 Cantonese) children in various tasks. Although the Cantonese-speaking children generally outperformed the English ones, Dowker and Li did not always find specific language effects that could be attributed to the greater transparency of the Chinese counting system and suggest that besides linguistic differences, cultural differences also deserve thorough consideration. Heubner et al. show that congruency effects observed in multi-digit number processing cannot only be reduced to magnitude-related influences. They also show that psychologically parity is not just an exact categorical attribute as it is mathematically. Multiple numerical and non-numerical factors influence parity judgments of two-digit numbers, and some of these influences are language specific as well.

EXTENSIONS—EXACT SNAs

This SNA category was also thoroughly investigated in the past and is very well-covered in the Research Topic. Studies on this SNA type fall into three groups. Firstly, some studies

have aimed to investigate the nature of this SNA type by proposing new experimental paradigms and developing new theories. Thompson et al. show that finding a specific page in a book correlates with number line estimation scores (even after controlling other variables), and this new task is postulated to tap into the same processes as the traditional number line estimation task (Siegler, 2009). Van't Noordende et al. show that children dynamically change strategies employed when solving the number line task. Secondly, Dowker and Li also utilized the number line estimation task, and show that although Chinese children outperform their English peers in more language-loaded task, the performance on the number line estimation task does not differ between the groups, suggesting less reliance of this SNA on language (see Helmreich et al., 2011 for opposite effects). Thirdly, Sella et al. as well as Opfer et al. focus on the functional role of this SNA type. Sella et al. show that preschoolers' ability to compare Arabic numbers (but not number words) relates to both ordinal and cardinal knowledge as well as the ability to order numbers in space. In the case of number words, magnitude comparison depended only on ordinal knowledge. Opfer et al. show that accuracy of linear mappings of numbers in the number line estimation task relates to multiple measures of memory for numbers and thus may provide a cornerstone or helpful tool for number memory.

All of these dissociations and impacts observed within this SNA type call for more fine-grained analyses and careful generalizations of conclusions.

EXTENSIONS—APPROXIMATE SNAs

Papers referring to this SNA category in the present Research Topic were mostly investigating the nature and structure of this SNA type, but also employed novel paradigms. Kucian et al. argue for the presence of a general magnitude system responsible for both spatial and numerical processing and attempt to observe its developmental trajectory. Their empirical evidence suggests that the structure of such a magnitude system is rather complex and consists of dissociated but related representations of continuous and discrete magnitudes. Rugani et al. provide evidence that numerical primes not only influence reaction time patterns but also affect motor execution, especially the reaching component of the movement. In sum, evidence coming from new task types supports SNA construct and criterion validity.

NOT IN TAXONOMY

Many, but not all the papers in this volume fit into the taxonomy. Nevertheless, they provide important insights into our understanding of human number processing. Pixner et al. investigate the development of the understanding of the cardinality of small numbers and zero. Zero is a special number and its role is hard to overestimate (see Nuerk et al., 2004; Nieder, 2016). Thus, it is crucial to understand the developmental trajectories of the understanding of the zero concept. Eventually, in the opinion piece, Fischer et al. provides multiple arguments for why numbers are embodied concepts. While it plays a crucial

role in some extension-exact SNAs (especially embodied number line trainings, Dackermann et al., 2017, for a review), its role for some directional SNAs, and especially place value-processing might be more limited.

WHERE DO WE STAND: CONCLUSIONS AND LIMITATIONS

Where are we with the SNA research? What are SNAs and what is their function in math skills and number processing? Summarizing existing studies, especially considering papers published within this Research Topic, we propose the following agenda for future SNA research.

Firstly, at least for now, we need to suspend our hopes about making conclusions as to what SNAs are in general since they seem to be a too broad range of phenomena and underlying representations. One can characterize some common features of some but not all SNA types. These common features may partially overlap. However, for the moment, the conclusions about SNAs, and their role need to be limited to specific SNA instances.

Secondly, both previous research (see Cipora et al., 2018a) and papers published in this Research Topic (e.g., Lohmann et al.; Bender et al.) show that SNAs are not stable or fixed, but they can be influenced by factors such as language, culture or even simple experimental manipulation. These inter- and intra-individual differences need to be taken into account when the functional role of SNAs is investigated.

We wish to summarize with the claim that space is an extremely important tool for many aspects of number processing

and representation. However, not all SNAs (e.g., Directional Implicit Cardinality SNAs) might be equally important for math skills. We still need to learn when space is important for math, but also when math is important for space: after all, in most of the studies we are looking at correlations. The example in which math skills influence strategies and also performance in spatial tasks has already been demonstrated (see Van 't Noordende et al.). For future investigation of the reciprocal interaction between SNA and math skills, it is crucial to specify both the SNAs and math skill under scrutiny, and carefully generalize the conclusions to other SNA types.

AUTHOR CONTRIBUTIONS

KC and H-CN wrote the first draft, which was then read, commented, and approved by FD and MH. All authors approved the final version of the manuscript.

FUNDING

KC and H-CN were supported by the DFG (NU 265/3-1) grant to H-CN. MH was supported by the National Science Centre, Poland (2014/15/G/HS6/04604) grant.

ACKNOWLEDGMENTS

We thank Zoë Kirste for language proofreading of the manuscript.

REFERENCES

- Cipora, K., He, Y., and Nuerk, H.-C. (2019). The SNARC effect and math skill: why should they be related? A model framework. *PsyArxiv Preprints*. doi: 10.31234/osf.io/aut45
- Cipora, K., Patro, K., and Nuerk, H.-C. (2015). Are Spatial-Numerical Associations a cornerstone for arithmetic learning? The lack of genuine correlations suggests No. *Mind Brain Educ.* 9, 190–206. doi: 10.1111/mbe.12093
- Cipora, K., Patro, K., and Nuerk, H.-C. (2018a). "Situational influences on spatial-numerical associations," in *Spatial Biases in Perception and Cognition*, ed T. Hubbard (Cambridge: Cambridge University Press), 41–49.
- Cipora, K., Schroeder, P. A., and Nuerk, H.-C. (2018b). "On the multitude of mathematics skills – Spatial-Numerical Associations and geometry skill?" in *Visualizing Mathematics: The Role of Spatial Reasoning in Mathematical Thought*, eds K. S. Mix and M. Battista (Cham: Springer), 361–370. doi: 10.1007/978-3-319-98767-5_18
- Cipora, K., Schroeder, P. A., Soltanlou, M., and Nuerk, H.-C. (2018c). "More space, better mathematics: Is space a powerful tool or a cornerstone for understanding arithmetic?" in *Visualizing Mathematics: The Role of Spatial Reasoning in Mathematical Thought*, eds K. S. Mix and M. T. Battista (Cham: Springer), 77–116.
- Dackermann, T., Fischer, U., Nuerk, H.-C., Cress, U., and Moeller, K. (2017). Applying embodied cognition: from useful interventions and their theoretical underpinnings to practical applications. *ZDM* 49, 545–557. doi: 10.1007/s11858-017-0850-z
- Dehaene, S., Bossini, S., and Giraux, P. (1993). The mental representation of parity and number magnitude. *J. Exp. Psychol. General* 122, 371–396.
- Di Giorgio, E., Lunghi, M., Rugani, R., Regolin, L., Dalla Barba, B., Vallortigara, G., et al. (2019). A mental number line in human newborns. *Dev. Sci.* 22:e12801. doi: 10.1111/desc.12801
- Gevers, W., Reynvoet, B., and Fias, W. (2003). The mental representation of ordinal sequences is spatially organized. *Cognition* 87, B87–B95. doi: 10.1016/s0010-0277(02)00234-2
- Helmreich, I., Zuber, J., Pixner, S., Kaufmann, L., Nuerk, H.-C., and Moeller, K. (2011). Language effects on children's nonverbal number line estimations. *J. Cross Cultural Psychol.* 42, 598–613. doi: 10.1177/0022022111406026
- Knops, A., Thirion, B., Hubbard, E. M., Michel, V., and Dehaene, S. (2009). Recruitment of an area involved in eye movements during mental arithmetic. *Science* 324, 1583–1585. doi: 10.1126/science.1171599
- Mix, K. S., Levine, S. C., Cheng, Y.-L., Young, C., Hambrick, D. Z., Ping, R., et al. (2016). Separate but correlated: the latent structure of space and mathematics across development. *J. Exp. Psychol. General* 145, 1206–1227. doi: http://dx.doi.org/10.1037/xge0000182
- Moeller, K., Shaki, S., Göbel, S. M., and Nuerk, H.-C. (2015). Language influences number processing - a quadrilingual study. *Cognition* 136, 150–155. doi: 10.1016/j.cognition.2014.11.003
- Nieder, A. (2016). Representing something out of nothing: the dawning of zero. *Trends Cogn. Sci.* 20, 830–842. doi: 10.1016/j.tics.2016.08.008
- Nuerk, H.-C., Iversen, W., and Willmes, K. (2004). Notational modulation of the SNARC and the MARC (linguistic markedness of response codes) effect. *Q. J. Exp. Psychol. Sect. A Human Exp. Psychol.* 57, 835–863. doi: 10.1080/02724980343000512
- Nuerk, H.-C., Weger, U., and Willmes, K. (2005a). Language effects in magnitude comparison: small, but not irrelevant. *Brain Lang.* 92, 262–277. doi: 10.1016/j.bandl.2004.06.107
- Nuerk, H.-C., Wood, G., and Willmes, K. (2005b). The universal SNARC effect: the association between number magnitude and space is amodal. *Exp. Psychol.* 52, 187–194. doi: 10.1027/1618-3169.52.3.187

- Patro, K., Fischer, U., Nuerk, H.-C., and Cress, U. (2016a). How to rapidly construct a spatial-numerical representation in preliterate children (at least temporarily). *Dev. Sci.* 19, 126–144. doi: 10.1111/desc.12296
- Patro, K., and Nuerk, H.-C. (2017). Limitations of trans-species inferences: the case of spatial-numerical associations in chicks and humans. *Cogn. Sci.* 41, 2267–2274. doi: 10.1111/cogs.12432
- Patro, K., Nuerk, H.-C., and Cress, U. (2016b). Mental number line in the preliterate brain: the role of early directional experiences. *Child Dev. Perspect.* 10, 172–177. doi: 10.1111/cdep.12179
- Patro, K., Nuerk, H.-C., Cress, U., and Haman, M. (2014). How number-space relationships are assessed before formal schooling: a taxonomy proposal. *Front. Psychol.* 5:419. doi: 10.3389/fpsyg.2014.00419
- Pinhas, M., and Fischer, M. H. (2008). Mental movements without magnitude? A study of spatial biases in symbolic arithmetic. *Cognition* 109, 408–415. doi: 10.1016/j.cognition.2008.09.003
- Rugani, R., Vallortigara, G., Priftis, K., and Regolin, L. (2015). Number-space mapping in the newborn chick resembles humans' mental number line. *Science* 347, 534–536. doi: 10.1126/science.aaa1379
- Shaki, S., Fischer, M. H., and Petrusic, W. M. (2009). Reading habits for both words and numbers contribute to the SNARC effect. *Psychon. Bull. Rev.* 16, 328–331. doi: 10.3758/PBR.16.2.328
- Siegler, R. S. (2009). Improving the numerical understanding of children from low-income families. *Child Dev. Perspect.* 3, 118–124. doi: 10.1111/j.1750-8606.2009.00090.x
- van Dijck, J.-P., and Fias, W. (2011). A working memory account for spatial-numerical associations. *Cognition* 119, 114–119. doi: 10.1016/j.cognition.2010.12.013
- Wood, G., Willmes, K., Nuerk, H.-C., and Fischer, M. H. (2008). On the cognitive link between space and number: a meta-analysis of the SNARC effect. *Psychol. Sci. Q.* 4, 489–525. Available online at: <https://psycnet.apa.org/record/2009-00781-003>

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Cipora, Haman, Domahs and Nuerk. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Developmental Dyscalculia and Automatic Magnitudes Processing: Investigating Interference Effects between Area and Perimeter

Hili Eidlin-Levy* and Orly Rubinsten

Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, Department of Learning Disabilities, University of Haifa, Haifa, Israel

OPEN ACCESS

Edited by:

Frank Domahs,
Philipps University of Marburg,
Germany

Reviewed by:

Ursula Fischer,
University of Regensburg, Germany
Annemie Desoete,
Ghent University, Belgium

*Correspondence:

Hili Eidlin-Levy
hileidlin@gmail.com

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 24 July 2017

Accepted: 04 December 2017

Published: 21 December 2017

Citation:

Eidlin-Levy H and Rubinsten O
(2017) Developmental Dyscalculia
and Automatic Magnitudes
Processing: Investigating Interference
Effects between Area and Perimeter.
Front. Psychol. 8:2206.
doi: 10.3389/fpsyg.2017.02206

The relationship between numbers and other magnitudes has been extensively investigated in the scientific literature. Here, the objectives were to examine whether two continuous magnitudes, area and perimeter, are automatically processed and whether adults with developmental dyscalculia (DD) are deficient in their ability to automatically process one or both of these magnitudes. Fifty-seven students (30 with DD and 27 with typical development) performed a novel Stroop-like task requiring estimation of one aspect (area or perimeter) while ignoring the other. In order to track possible changes in automaticity due to practice, we measured performance after initial and continuous exposure to stimuli. Similar to previous findings, current results show a significant group \times congruency interaction, evident beyond exposure level or magnitude type. That is, the DD group systematically showed larger Stroop effects. However, analysis of each exposure period showed that during initial exposure to stimuli the DD group showed larger Stroop effects in the perimeter and not in the area task. In contrast, during continuous exposure to stimuli no triple interaction was evident. It is concluded that both magnitudes are automatically processed. Nevertheless, individuals with DD are deficient in inhibiting irrelevant magnitude information in general and, specifically, struggle to inhibit salient area information after initial exposure to a perimeter comparison task. Accordingly, the findings support the assumption that DD involves a deficiency in multiple cognitive components, which include domain-specific and domain-general cognitive functions.

Keywords: developmental dyscalculia, magnitude processing, geometric processing Stroop task, inhibition processing

INTRODUCTION

In the past few years, the relationship between numbers and other magnitudes, such as space and time, has evoked a great deal of interest (Walsh, 2003; Dehaene and Brannon, 2010; Gebuis and Reynvoet, 2012a; Newcombe, 2014; Leibovich et al., 2016; McCaskey et al., 2017). Walsh (2003) was the first to suggest the existence of a common processing mechanism for time, space, and quantity, and he established a Theory of Magnitude (ATOM). Evolutionarily, such a mechanism enabled simultaneous processing of numerical, temporal, and spatial features of the world in order to produce adaptive reactions. Following Walsh's theory, a wealth of knowledge appeared

with regards to magnitude and numerical associations (Pinel et al., 2004; Dakin et al., 2011; Ren et al., 2011; Stoianov and Zorzi, 2012; Tibber et al., 2012; Leibovich et al., 2016; McCaskey et al., 2017). Specifically, recent research implies the existence of a magnitude estimation mechanism representing both numerical and other continuous magnitudes, including area, perimeter, length, volume, time, etc. (Lourenco and Longo, 2010; Dakin et al., 2011; Gebuis and Reynvoet, 2012b; Leibovich and Henik, 2013, 2014; Newcombe, 2014; Leibovich et al., 2016; Lourenco and Bonny, 2016).

Accumulating evidence shows that magnitude information is probably more perceptually salient than numerical information, and thus may have developed earlier (Feigenson et al., 2002; Gebuis and Reynvoet, 2012a; Odic et al., 2012). As an example, infant research shows that continuous magnitudes, such as the contour length (perimeter) or size (area) of figures, are more salient to infants than the number of items (Clearfield and Mix, 1999; Feigenson et al., 2002). Despite the fact that continuous magnitude processing seems to be as important as numerical processing for daily performance (Gebuis and Reynvoet, 2012b; Leibovich and Henik, 2013, 2014), comparisons between different types of continuous magnitudes (e.g., area and perimeter) have received much less scientific attention.

The current study aims to deepen the knowledge about continuous magnitude processing and has three main purposes. The first purpose is to investigate automatic processing (namely, processed spontaneously and with no need for monitoring; Tzelgov, 1997) of two continuous magnitudes, area and perimeter. The second purpose is to explore general interactions between numerical cognition and magnitude processing by inclusion of a specific clinical population – participants with deficient numerical abilities (developmental dyscalculia or DD). The third purpose is to investigate whether the DD group benefits from continuous exposure to stimuli as much as the control group.

Area and Perimeter Processing

As mentioned, the study aims to investigate the automaticity of area and perimeter magnitudes. The automaticity of magnitude processing is investigated here via a Stroop-like task. Generally, in Stroop tasks it is typically found that participants cannot ignore irrelevant dimensions of the task (e.g., the physical size of a digit), which are processed involuntarily and interfere with the processing of the relevant dimension (e.g., the actual quantity that the digit represents; Henik and Tzelgov, 1982). Here, we adopted a Stroop-like comparison task created by Babai et al. (2006), which compared the automaticity of area vs. perimeter processing. Previous findings demonstrated that in a perimeter comparison task, the figure's area was processed involuntarily and thus affected perimeter judgment. In the *congruent condition*, the figure with the greater perimeter also had a larger area, while in the *incongruent condition*, the perimeters of the two figures were equal, while one of them had a larger area. Results indicated that participants showed a clear *Stroop effect*, namely they responded faster to congruent than to incongruent trials (Stavy et al., 2006; Stavy and Babai, 2008, 2010). Thus, the irrelevant area aspect might be automatically processed and, despite being irrelevant

to the task, interfere with perimeter processing. Notably, the opposite task, that is when participants were asked to determine which figure has the largest area (and to ignore the irrelevant perimeter) was examined as well. Results showed that of the two tasks, area comparisons were significantly faster than perimeter comparisons. However, unlike the perimeter comparison task (where area is irrelevant to the task), in the area comparison task (where perimeter is irrelevant to the task) no differences were found between congruent and incongruent trials (Babai et al., 2006). The researchers inferred that area is more salient and more rapidly processed (i.e., more automatic) than perimeter.

To this end, it is not clear whether one magnitude (area) can be more perceptually salient than the other (perimeter). According to general magnitude mechanism theories, both magnitudes should be involuntarily processed and hence are supposed to interfere with each other to a similar degree (Newcombe, 2014; Leibovich et al., 2016). On the other hand one magnitude, in this case area (as found in Babai et al., 2006), can interfere with the processing of another magnitude (perimeter), but not vice versa. A possible solution for this contradiction is that magnitude saliency differs due to task demands (Spelke et al., 2010). Spelke et al. (2010) identified two core systems used in geometric processing: the first is important for navigation at large – scale surfaces (Cheng and Newcombe, 2005) while the second system is used for small object recognition, such as the figures in Babai et al.'s (2006) task. Each system relates to different neural and cognitive processing, with limited transference between them (Derdikman and Moser, 2010; Huang and Spelke, 2015). Accordingly, a particular magnitude can be automatically processed by one of the core systems but not by the other. For instance, the processing of a surface's layout (or, in other words, *perimeter*) is highly crucial for successful navigation (Hermer and Spelke, 1994; Hermer-Vazquez et al., 1999; Lee et al., 2012) but not as crucial for recognition of 2D figures (Lee and Spelke, 2011).

Developmental Dyscalculia and Magnitude Processing

The current research also aims to expand the knowledge about the interactions between numerical cognition and magnitude processing by including a specific clinical population – participants with deficient numerical abilities (DD). DD is a specific deficit of numerical and mathematical abilities, with a neuro-anatomical source (Rotzer et al., 2008; Kaufmann et al., 2013), affecting about 3.6–6.5% of the population (Geary, 1993; Shalev and Gross-Tsur, 2001; Butterworth, 2005; Reigosa-Crespo et al., 2012). Individuals with DD fail to master common numerical and arithmetical skills, such as numerical comparisons (Rubinsten and Henik, 2005; Mussolin et al., 2010a), arithmetical fact retrieval (Mazzocco et al., 2008), and procedural knowledge (Desoete et al., 2009). Currently, there is a debate about the cognitive mechanisms that underlie numerical difficulties, and DD seems to be a heterogeneous disorder (Rubinsten and Henik, 2009; Kaufmann et al., 2013; Träff et al., 2016). One main explanation for the heterogeneity of DD is that numerical ability cannot be described as a single cognitive mechanism and arithmetic competence relies on domain-specific as well as domain-general skills (Rubinsten and Sury, 2011;

Kaufmann et al., 2013; Szűcs and Goswami, 2013; Bugden and Ansari, 2015; Huber et al., 2015; Noël et al., 2016). From one perspective, individuals with DD may have a domain-specific deficient “number sense,” manifested by difficulties with representing and manipulating all kinds of numerical notations: non-symbolic – such as dot arrays (Price et al., 2007; Mussolin et al., 2010b), or symbolic, namely Arabic numerals and number words (Rubinsten and Henik, 2005, 2006; Mussolin et al., 2010a). Concurrent with the general magnitude processing mechanism theory, they may also ineffectively process continuous magnitudes and have deficient “magnitude sense” (Leibovich et al., 2016). It is worth noticing though, that in a recent study adolescents with DD showed similar performance, on both behavioral and neuronal levels, as typical developing peers when performing non-symbolic numerical comparison tasks (McCaskey et al., 2017). However, participants with typical development activated domain-specific magnitude related areas while performing the task, while participants with DD activated domain-general frontal areas, related to inhibition and working memory. Accordingly, the authors inferred that domain-general deficits could also account for the development of DD.

Indeed, the role of inhibition, a domain-general mechanism, in intact and deficient numerical processing has also received scientific attention (Ashkenazi et al., 2009; Ashkenazi and Henik, 2010; Wang et al., 2012; Szucs et al., 2013a; Bugden and Ansari, 2015; Noël et al., 2016). From this perspective, individuals with DD show deficient performance on numerical tasks due to failure to inhibit irrelevant magnitude information, such as the overall area of dot arrays in a non-symbolic comparison task (Bugden and Ansari, 2015) or physical size in the numerical Stroop task (Szucs et al., 2013a). It is important to notice that these findings remained consistent after controlling for comorbidity (Wang et al., 2012).

The Current Research

The discussion of a general magnitude processing mechanism as a basis for numerical cognition development seems to be more relevant than ever (Gebuis and Reynvoet, 2012b; Leibovich and Henik, 2014; Newcombe, 2014; Lourenco and Bonny, 2016; Leibovich et al., 2016). However, there exists no extensive work dealing with intact and deficient processing of this system (Skagerlund and Träff, 2014; McCaskey et al., 2017). Moreover, it is necessary to differentiate between “pure,” domain-specific, and domain-general mechanisms when performing magnitude comparison tasks, in order to define the source of DD difficulties.

Hence, we compared participants with DD and typically developing participants while performing a Stroop-like task (adopted from Babai et al., 2006). We expected to find group differences, as measured by Stroop effects (incongruent minus congruent trials) in both area and perimeter comparison tasks. According to domain-specific magnitude processing deficits (Newcombe, 2014; Leibovich et al., 2016), the DD group is expected to show smaller Stroop effects implying that they do not process the irrelevant magnitude (similar to the numerical Stroop task, Rubinsten and Henik, 2005). On the other hand, domain-general deficits (Wang et al., 2012; Szucs et al., 2013a; Bugden and Ansari, 2015) should result in larger Stroop effects in the DD

group, due to a deficit in the ability to ignore irrelevant magnitude information.

Another interesting specific question is whether participants with DD will show similar deficient processing on both the area and perimeter tasks. If the area component is indeed more perceptually salient (Babai et al., 2006), the DD group should show larger Stroop effects on the perimeter task (in which area is irrelevant to the task), namely they should find it harder to ignore the irrelevant area aspect.

The cognitive method as well as the statistical analysis of the current study enabled us not only to study the differences between automatic processing of area vs. perimeter (i.e., investigating magnitude sense in DD), but also to investigate whether DD participants indeed perform poorly or differently on continuous magnitude processing tasks in initial vs. proficiency stages of learning (i.e., to investigate learning functions in DD). Earlier studies showed that even a small number of rehearsals of numerical problems led to automatic processing and to changes in brain functions (Ischebeck et al., 2007; Aubin et al., 2016). For instance, Ischebeck et al. (2007) found that very short training (eight repetitions) in multiplication problems led to a decrease in the activity of fronto-parietal brain areas related to calculation and numerical processing (Menon et al., 2000; Dehaene et al., 2003). On the other hand, the training also resulted in increased activity in temporo-parietal regions known to be involved in arithmetic fact retrieval (Dehaene et al., 2003). Recently, it was proposed that DDs’ deficits in inhibition of irrelevant numerical information can also represent difficulties with consolidating learned information and with performing the shift from initial computing based processing to automatic retrieval based processing (Ischebeck et al., 2007; Aubin et al., 2016). Indeed, Aubin et al. (2016) suggested that people with DD may be less able to consolidate a numerical task within the frontal-parietal region and must instead rely on their working memory. Consequently, there would be limited progression to recalling numerical information and a continued dependence on working memory. Accordingly, it is predicted here that people with DD may need more rehearsals in order to attain proficiency in performing magnitude tasks. Such difficulty with consolidating learned knowledge and with using advanced retrieval strategies will produce consistent group differences predicted to be prominent in the perimeter task, evident even after continuous exposure to stimuli.

MATERIALS AND METHODS

Participants

Fifty-seven adults, 27 typically developing (i.e., control group; including 9 males, 18 females; mean age = 24.92 years, $SD = 2.67$ years) and 30 with DD (2 males, 28 females; mean age = 24.43 years, $SD = 2.74$ years) participated in the study. All participants were university students and had successfully completed math matriculation exams. Participants were recruited by advertisements distributed on campus and gave written consent to participate in the experiment. Some of them were paid about USD15 for their participation, while other received

credit points for academic courses. The study was carried out in accordance with the recommendations of the ethics committee of the University of Haifa with written informed consent from all participants. All participants gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the ethics committee of the University of Haifa (No. 108/09).

Classification and Assessment Criteria

Participants were assigned to either a control or a DD group, using the “Israeli learning function diagnosis system” (also titled “MATAT” in Hebrew) for high school and higher education students (National Institute for Testing and Evaluation). This diagnostic tool is composed of a set of standardized computerized tests and questionnaires intended for diagnosing learning disabilities in high school and higher education students. All tests and questionnaires used are nationally normalized. All participants performed numerical tasks (simple calculation and procedural knowledge calculation) and reading related tasks (text reading, phoneme omission, and rapid naming). They also answered a questionnaire (based on the DSM) regarding their attention ability, and performed Raven’s Standard Progressive Matrices (SPM) test (Raven et al., 1995) in order to rule out non-verbal mental disabilities.

Participants were defined as having DD if their scores in either RT or accuracy (ACC) measures on the simple calculation and procedural knowledge tests were worse than mean -1 SD and their scores in the reading and attention tests were more than $+1$ SD above the mean. Participants in the control group scored better than mean -1 SD in numerical, reading, and attention tests. All participants also scored above the 25th percentile in the SPM test. Independent t -tests were conducted on the different test results. The two groups differed significantly in both numerical tests (for mean test results and p -values of independent t -tests see **Table 1**). As significant group differences were also evident in reading related tasks, group results did not exceed 1SD lower than the norm score, hence it can be assumed that all participants were good readers.

The Experimental Tasks – Area and Perimeter Tasks

The experiment was run on a PC using E-Prime 2.0 software, and contained two tasks, based on the comparison task of Babai et al. (2006). In each task, participants were requested to relate to a specific aspect of the stimuli (i.e., area or perimeter). Participants were presented with two polygons in each trial, and were asked to decide which polygon had the largest amount of one of the magnitude aspects (area/perimeter). Each block contained a similar number of congruent trials (where both area and perimeter are increased between stimuli) and incongruent trials (where one aspect is increased while the other is reduced between stimuli). Semi-neutral trials (where one aspect is equal and the other differs between stimuli) and equal trials (where both aspects are equal between stimuli) were added as fillers.

Stimuli

Each stimulus was composed of two figures – a basic rectangle or a polygon derived from the basic rectangle by adding or subtracting one or two squares (the size of each square was $1/12$ of the basic rectangle size). Pairing of polygons was based on their protrusion direction (into the basic rectangle, see **Figure 1B**, or protruding out of it, see **Figure 1A**) and congruency rules (described above). Each polygon was ascribed to two different stimuli in order to avoid visual bias. Moreover, in order to avoid visual bias and learning, each stimulus was presented eight times, four times on the left side of the fixation point and four times on the right side of the fixation point. Each stimulus was also presented twice in four directions – two on the vertical axis – original (turning up) or rotated by 180° (turning down) and two on the horizontal axis – rotated by 90° to the left (turning left) or to the right (turning right).

The task contained two blocks, one with filled figures and the other containing framed only figures, in order to avoid visual bias of one aspect (Stavy and Babai, 2008). Each block contained 32 “initial exposure” trials (following Ischebeck et al., 2007), representing the *practice* phase. We aimed to capture the reaction to unfamiliar stimuli – eight congruent trials (both area and perimeter were larger in one of the polygons), eight incongruent trials (where the area was larger for one polygon while the perimeter was larger in the other), and 16 filler trials (in which one or both aspects remained equal). After succeeding on more than 80% of the practice trials (in the first or second presentation), participants were able to continue to the continuous exposure, or *proficiency* phase, representing automatization levels. The proficiency phase contained 128 trials – 32 congruent trials (both area and perimeter were larger in one of the polygons), 32 incongruent trials (where the area was larger for one polygon while the perimeter was larger in the other), and 64 filler trials (in which one or both aspects remained equal). The filler trials enabled sufficient presentation of each stimulus and contained 32 equal trials, with the same polygons presented at different rotations, and 32 semi-neutral trials, where one aspect (area or perimeter) remained constant between polygons while the other differed. Since there were two experimental tasks – area and perimeter comparison – each experimental condition, congruent and incongruent, was represented in total by 32 trials for the practice phase (8×2 blocks, filled or framed $\times 2$ tasks – area and perimeter) and 128 trials for the proficiency phase (32×2 blocks, filled or framed $\times 2$ tasks – area and perimeter). For illustration of the experimental blocks see **Figure 2**.

Procedure

Participants were seated about 60 cm from the computer screen. The participants’ goal was to decide, in two separate tasks, which polygon has the largest area (area task) or perimeter (perimeter task). Then, they were asked to convey their answer as quickly and accurately as possible by pressing one of three marked keys on a response box (right key for right polygon, left key for left polygon, central key for equal area or perimeter). Each participant was presented with

TABLE 1 | Mean standard scores and group differences in the screening tests of the two groups.

Test		Control		DD		t-test	
Mean scores		ACC	RT	ACC	RT	ACC	RT
Text reading		0.56	0.42	0.48	−0.31	0.37	2.78**
Phoneme omission		0.46	0.65	−0.05	0.01	2.34*	2.53*
Rapid naming	Objects		0.14		−0.32		1.94
	Letters		0.64		0.33		1.21
	Numbers		0.42		−0.15		2.08*
Questionnaire	Attention	0.09		0.12			−0.37
	Impulsiveness and hyperactivity	0.00		0.86			−1.57
Simple calculation		0.61	0.49	−1.3	−1.37	4.79**	5.93**
Procedural knowledge		0.46	0.59	−1.55	−1.26	6.05**	6.92**
SPM (percentiles)		58.4		52.43		1.25	

ACC = standard scores for accuracy rates; RT = standard scores for reaction times; SPM = Standard Progressive Matrices (Raven et al., 1995); Sig. = independent t-test significance between control and DD groups. Significance values: * $p < 0.05$; ** $p < 0.01$.

four blocks (two area blocks – filled and framed, and two perimeter blocks – filled and framed). Half the participants began with the area task (25% of participants began with a filled figures block and the rest with a framed figures block) and half with the perimeter task (25% began with a filled figures block and the rest with a framed figures block). Tasks and block performance order were counter balanced and were determined in advance according to participants' serial numbers.

The experiment included three breaks in each block, which were terminated when participants pressed a relevant key, as well as a break of a few minutes between the sections. The stimuli in each trial began with a fixation point (small white asterisk), which appeared for 300 ms and was followed by an empty black screen for 500 ms. Then the sample polygons appeared and remained in view until the participant pressed a key but no longer than 5,000 ms. The next trial began with the fixation point. Overall, the experimental tasks took about 1 h.

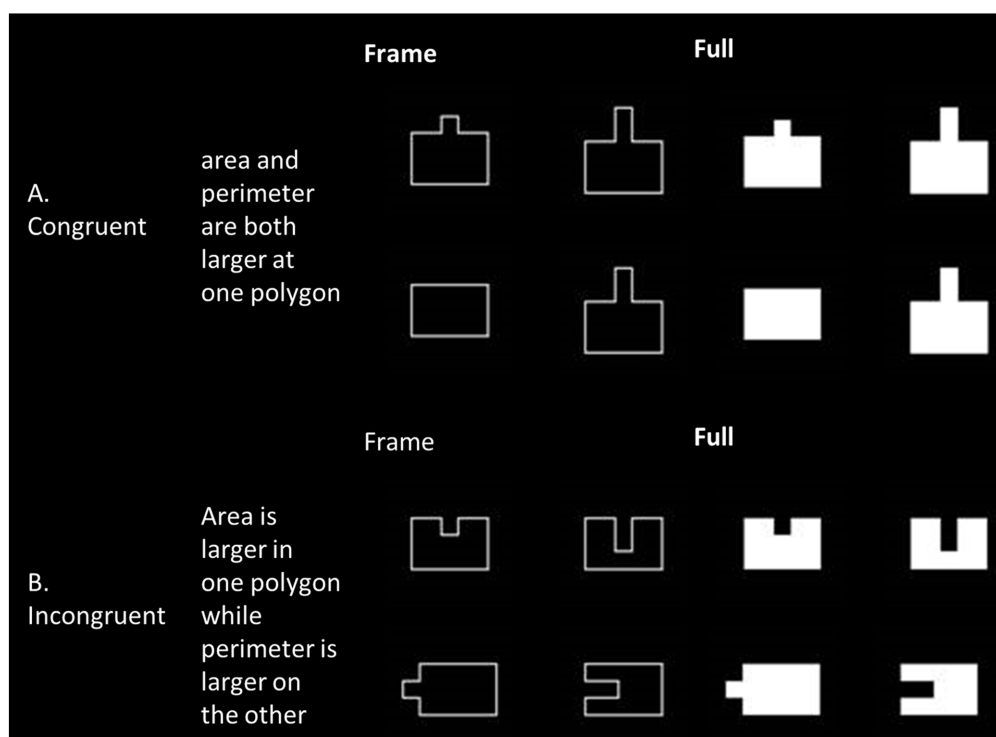


FIGURE 1 | Examples of framed and filled stimuli, arranged by congruency. **(A)** Examples of congruent stimuli, where both area and perimeter are increased in both polygons. **(B)** Examples of incongruent stimuli, where a different aspect is increased in each polygon.

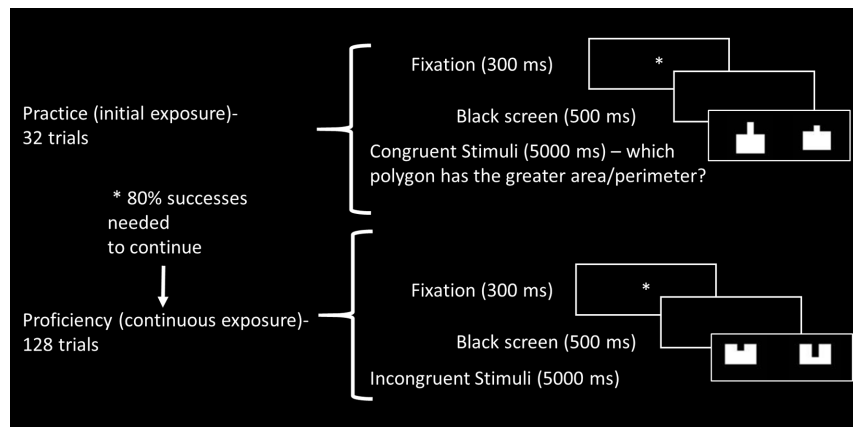


FIGURE 2 | Illustration of experimental block, including initial (practice) and continuous exposure (proficiency) trials.

Data Analysis

Error Rates Analysis

Mean error rates (percentage of incorrect trials) were calculated for each participant for the practice (only the first practice was calculated for participants who failed to succeed in over 80% of the practice trials) and proficiency phases separately. Then, three-way repeated measures ANOVAs were used, with group (i.e., Control or DD) as the between-subject factor, and congruency (i.e., congruent or incongruent) and task (i.e., Area or Perimeter) as within-subject factors on the error rates of practice and proficiency trials. All of the following F -statistics were adjusted by the Greenhouse-Geisser correction.

In order to define the triple group \times task \times congruency interaction, Stroop effects of the error rates were calculated (by subtracting error rates of incongruent trials from error rates of congruent trials) for each group, in Area and Perimeter tasks separately. In order to test for a continuous exposure effect, we compared the Stroop effects of practice and proficiency phases of each task for each group separately, by using paired t -tests.

Reaction Times (RT) Analysis

Reaction times analysis was similar to error rates analysis. Mean RTs (in ms) were calculated for each participant for practice and proficiency phases separately. Then, three-way repeated measures ANOVAs were used, with group as the between-subject factor and congruency and task as within-subject factors on the RTs of practice and proficiency trials. A continuous exposure effect was observed by comparing the Stroop effects of the practice and proficiency phases of each task for each group separately, using paired t -tests.

Other Visual Features

We used repeated measures ANOVAs, similar to those described above, in order to investigate whether other visual features manipulated (i.e., figure filling, protrusion, and direction – see **Figure 1**) were possible confounders of Area and Perimeter processing.

Gender Differences

Gender differences were tested using independent t -tests.

RESULTS

Gender Differences

The research sample contained a larger number of females. Accordingly, we tested gender differences (mean scores and t -tests are presented in **Table 2**). Females and males showed similar patterns across tasks and proficiency stages. Specifically, they showed similar Stroop effects in both tasks in both error rates and speed analyses.

Visual Features

Several visual features were manipulated in order to track possible confounding with the experimental variables. The filling effect (filled vs. framed figures) was not significant in the practice phase, neither for error rates [$F_{(1,55)} = 0.10, p = 0.746, \eta^2 = 0.002$] nor for RTs [$F_{(1,55)} = 0.39, p = 0.533, \eta^2 = 0.008$]. Moreover, there was no filling effect for error rates in the proficiency phase [$F_{(1,55)} = 0.11, p = 0.741, \eta^2 = 0.002$]. However, participants responded more slowly to framed figures ($M = 977.95, SD = 229.58$) than to filled figures ($M = 919.18, SD = 194.74$) in the proficiency phase [$F_{(1,55)} = 6.31, p = 0.015, \eta^2 = 0.103$]. Furthermore, in this phase, a filling \times congruency interaction was evident [$F_{(1,55)} = 5.28, p = 0.025, \eta^2 = 0.08$], such that the Stroop effect (incongruent minus congruent) was significantly greater [$F_{(1,56)} = 4.58, p = 0.037, \eta^2 = 0.07$] for framed figures ($M = 236.15, SD = 133.19$) than for full figures ($M = 196.77, SD = 106.72$).

We further analyzed the effect of continuous exposure for each filling type separately. This was done in order to figure out whether stimuli rehearsal led to changes in the automaticity of the filling effect. On the error analysis, participants made more errors (in percentage) in congruent trials than in incongruent trials in framed figures after initial exposure (M Stroop effect = $-0.05, SD = 0.07$) and this difference decreased after continuous exposure (M Stroop effect = $-0.00, SD = 0.02$) to stimuli

$[t_{(56)} = -4.06, p < 0.001, d = 1.085]$. For full figures, no difference between Stroop effects after initial ($M = -0.03, SD = 0.07$) or continuous exposure ($M = -0.01, SD = 0.04$) was evident [$t_{(56)} = -1.51, p = 0.137, d = -0.403$]. Reaction times analysis show that the Stroop effects (in ms) were larger after initial exposure and decreased after continuous exposure for both framed [initial exposure: $M = 326.46, SD = 201.91$; continuous exposure: $M = 243.28, SD = 134.87$; $t_{(56)} = 3.71, p = 0.001, d = 0.991$] and full figures [initial exposure: $M = 290.99, SD = 213.14$; continuous exposure: $M = 211.03, SD = 123.41$; $t_{(56)} = 2.86, p = 0.006, d = 0.764$].

There were no other interactions involving the filling component. No significant findings were found for error rates or RTs for other visual features manipulated, namely protrusion (protruding in or out – see **Figure 1**) and vertical or horizontal direction.

Error Rates Analysis

Practice Phase

The current analysis aims to define whether Area interferes with Perimeter processing and vice versa, among participants with intact and deficient numerical processing after initial exposure to non-familiar stimuli.

Mean error rates (in percentage) in area and perimeter tasks across different exposure phases of the two groups are presented in **Table 3**. Results revealed a significant effect of group [$F_{(1,55)} = 5.49, p = 0.023, \eta^2 = 0.091$], indicating that the DD group made more errors than the Control group. The main effect of congruency was also significant [$F_{(1,55)} = 39.81, p < 0.001, \eta^2 = 0.420$], as congruent trials had smaller error rates than incongruent trials. No main effect of task was evident [$F_{(1,55)} = 3.56, p = 0.064, \eta^2 = 0.061$], with similar error rates for both Area and Perimeter tasks.

However, the task \times congruency interaction was significant [$F_{(1,55)} = 22.72, p < 0.001, \eta^2 = 0.292$]. Further analysis revealed that incongruent trials were less accurate than congruent trials on the Perimeter task [$t_{(56)} = 6.7, p < 0.001, d = 1.79$]. No difference between error rates of congruent and incongruent trials was evident in the Area task [$t_{(56)} = 1.32, p = 0.192, d = 0.352$]. Importantly, no order effect was evident, as participants who started with the Perimeter task showed similar Stroop effects as participants starting with the Area task in both tasks [Area task:

$t_{(55)} = 0.767, p = 0.446, d = 0.206$; Perimeter task: $t_{(55)} = 0.004, p = 0.997, d = 0.001$].

Interestingly, and with high relevance to the current research questions, a triple interaction was evident [$F_{(1,55)} = 6.69, p = 0.012, \eta^2 = 0.109$], as described in **Figure 3**. Further analysis of this interaction revealed significant effects for group [$F_{(1,55)} = 4.73, p = 0.034, \eta^2 = 0.079$] and congruency [$F_{(1,55)} = 47.72, p < 0.001, \eta^2 = 0.454$] in the Perimeter task (in which Area is irrelevant to the task). Moreover, in this Perimeter task, the group \times congruency interaction reached significance as well [$F_{(1,55)} = 3.81, p = 0.050, \eta^2 = 0.065$]. The Stroop effect (incongruent minus congruent) of error rates was larger for the DD group ($M = 0.10, SD = 0.09$) than for the Control group ($M = 0.05, SD = 0.08$). In the Area task, no main effect of group [$F_{(1,55)} = 1.87, p = 0.176, \eta^2 = 0.033$] or congruency [$F_{(1,55)} = 2.04, p = 0.158, \eta^2 = 0.036$] was evident, nor group \times congruency interaction [$F_{(1,55)} = 2.8, p = 0.098, \eta^2 = 0.049$]. In other words, in the Area task, there was no meaningful Stroop effect of error rates, and this pattern was similar for both Controls ($M = 0.02, SD = 0.06$) and DD ($M = 0.00, SD = 0.06$). In conclusion, while the DD group showed larger Stroop effects in the perimeter task, no group difference was evident in the area task.

Proficiency Phase

Contrary to the former analysis, this analysis aims to find out whether magnitude interference and group differences are evident after continuous exposure, when participants are well familiar with the stimuli.

The main effect of task [$F_{(1,55)} = 5.80, p = 0.019, \eta^2 = 0.096$] was minor but significant. Importantly, no order effect was evident, as participants who started with the Perimeter task showed similar Stroop effects as participants starting with the Area task in both tasks [Area task: $t_{(55)} = 1.3, p = 0.196, d = 0.351$; Perimeter task: $t_{(55)} = -0.55, p = 0.582, d = -0.148$]. The main effect of congruency reached significance as well [$F_{(1,55)} = 18.63, p < 0.001, \eta^2 = 0.253$]. However, no main effect of group was evident [$F_{(1,55)} = 0.03, p = 0.852, \eta^2 = 0.001$], nor any interaction. This suggests that there was a typical Stroop effect, which was similar in pattern across groups (DD and Controls) and tasks (Area and Perimeter).

TABLE 2 | Mean scores and gender differences in area and perimeter tasks across different exposure phases.

	Task	Females <i>M</i> (<i>SD</i>)	Males <i>M</i> (<i>SD</i>)	<i>t</i> -test	<i>p</i> -value
Stroop effects of error rates (in percentage)	Practice phase- Area	0.01 (0.06)	0.01 (0.06)	0.468	0.642
	Practice phase- Perimeter	0.08 (0.09)	0.05 (0.07)	1.179	0.244
	Proficiency phase- Area	0.02 (0.04)	0.00 (0.03)	1.35	0.183
	Proficiency phase- Perimeter	0.01 (0.03)	0.01 (0.02)	-0.619	0.539
Stroop effects of reaction times (in ms)	Practice phase- Area	308.92 (217.90)	194.09 (205.24)	1.69	0.097
	Practice phase- Perimeter	378.85 (287.00)	266.23 (272.40)	1.25	0.214
	Proficiency phase- Area	227.66 (101.57)	270.58 (201.45)	-1.04	0.300
	Proficiency phase- Perimeter	242.73 (175.33)	193.76 (85.69)	0.969	0.337

Stroop effects = incongruent minus congruent trials. Sig. = independent t-test between females and males.

TABLE 3 | Mean error rates (in percentage) in area and perimeter tasks across different exposure phases of the two groups.

Group	Practice phase				Proficiency phase			
	Area		Perimeter		Area		Perimeter	
	Congruent <i>M</i> (<i>SD</i>)	Incongruent <i>M</i> (<i>SD</i>)	Congruent <i>M</i> (<i>SD</i>)	Incongruent <i>M</i> (<i>SD</i>)	Congruent <i>M</i> (<i>SD</i>)	Incongruent <i>M</i> (<i>SD</i>)	Congruent <i>M</i> (<i>SD</i>)	Incongruent <i>M</i> (<i>SD</i>)
Control	0.02 (0.03)	0.04 (0.05)	0.01 (0.02)	0.07 (0.07)	0.03 (0.03)	0.04 (0.03)	0.02 (0.02)	0.03 (0.03)
DD	0.05 (0.06)	0.05 (0.05)	0.01 (0.02)	0.12 (0.09)	0.03 (0.01)	0.04 (0.02)	0.01 (0.01)	0.03 (0.04)

Continuous Exposure Effect (i.e., The Statistical Differences between Stroop Effects in the Practice vs. Task Phases)

The effect of exposure was observed by comparing the Stroop effects in the practice and proficiency phases of each task for each group separately, using paired *t*-tests. This in order to figure out whether stimuli rehearsal led to changes in automaticity of one or both magnitudes among different participants.

As described in **Figure 3**, in the Perimeter task a significant difference was found between the Stroop effects of the practice and proficiency phases for the DD group [$t_{(29)} = 5.37$, $p < 0.001$, $d = 1.994$; practice phase: $M = 0.10$, $SD = 0.09$; proficiency phase: $M = 0.015$, $SD = 0.02$]. A similar significant difference was evident for the Control group [$t_{(26)} = 2.62$, $p = 0.015$, $d = 1.027$; practice phase: $M = 0.05$, $SD = 0.08$; proficiency phase: $M = 0.01$, $SD = 0.03$]. Accordingly, the initial Stroop effects declined after practice and this pattern was evident in both groups. No difference was evident in the Stroop effects in the Area task neither for the DD group [$t_{(29)} = -1.57$, $d = -0.583$, $p = 0.126$; practice phase: $M = 0.002$, $SD = 0.06$; proficiency phase: $M = 0.02$, $SD = 0.04$] nor for the Control group [$t_{(26)} = 1.03$, $p = 0.309$, $d = 0.403$; practice phase: $M = 0.02$, $SD = 0.06$; proficiency phase: $M = 0.01$, $SD = 0.04$].

Reaction Time Analysis (RT) Practice Phase

Mean reaction times (in ms) in area and perimeter tasks across different exposure phases of the two groups are presented in **Table 4**. When testing whether Area interferes with Perimeter processing and vice versa among participants with different numerical abilities after initial exposure to non-familiar stimuli, a main effect of congruency was found [$F_{(1,55)} = 215.54$, $p < 0.001$, $\eta^2 = 0.797$]. Accordingly, participants responded more slowly to incongruent trials than to congruent trials. The main effect of group was not significant [$F_{(1,55)} = 1.74$, $p = 0.191$, $\eta^2 = 0.031$], nor was the main effect of task [$F_{(1,55)} = 2.17$, $p = 0.146$, $\eta^2 = 0.038$].

However, and following the experimental hypothesis, the group \times congruency interaction was found to be significant [$F_{(1,55)} = 5.09$, $p = 0.028$, $\eta^2 = 0.085$]. The Stroop effect (incongruent minus congruent) was significantly larger [$F_{(1,55)} = 2.25$, $p = 0.028$] for the DD group ($M = 363.9$, $SD = 147.86$) than for the Control group ($M = 266.89$, $SD = 176.33$).

No triple interaction was found [$F_{(1,55)} = 0.87$, $p = 0.355$, $\eta^2 = 0.016$]. However, following scientific background and interest, we further analyzed the triple interaction (see **Figure 4**).

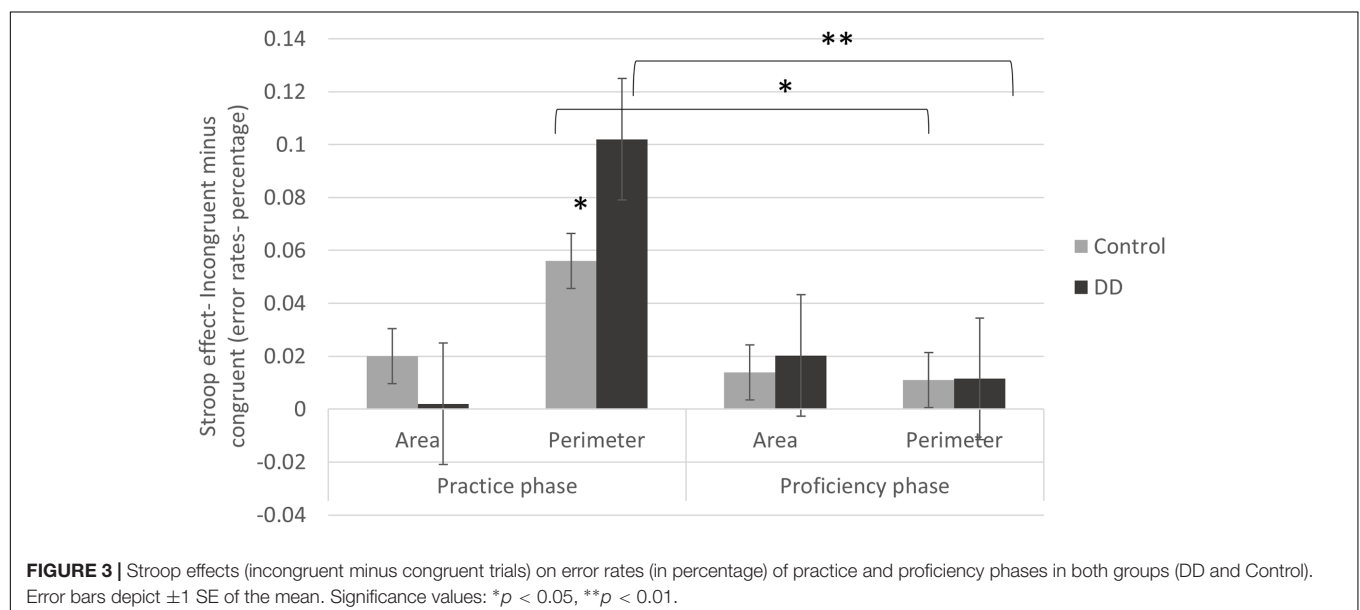


TABLE 4 | Mean reaction times (in ms) in area and perimeter tasks across different exposure phases of the two groups.

Group	Practice phase				Proficiency phase			
	Area		Perimeter		Area		Perimeter	
	Congruent <i>M</i> (<i>SD</i>)	Incongruent <i>M</i> (<i>SD</i>)	Congruent <i>M</i> (<i>SD</i>)	Incongruent <i>M</i> (<i>SD</i>)	Congruent <i>M</i> (<i>SD</i>)	Incongruent <i>M</i> (<i>SD</i>)	Congruent <i>M</i> (<i>SD</i>)	Incongruent <i>M</i> (<i>SD</i>)
Control	1087.50 (246.43)	1344.09 (213.55)	1144.53 (273.77)	1421.73 (396.78)	787.55 (135.92)	995.53 (200.28)	812.73 (137.41)	1008.33 (269.53)
DD	1153.47 (289.83)	1459.74 (267.62)	1142.79 (296.73)	1564.34 (366.68)	898.98 (181.51)	1162.95 (238.72)	896.42 (249.88)	1160.36 (281.49)

In the Perimeter task, a significant congruency effect was found [$F_{(1,55)} = 89.53$, $p < 0.001$, $\eta^2 = 0.619$], but no group effect [$F_{(1,55)} = 0.74$, $p = 0.391$, $\eta^2 = 0.013$]. Furthermore, a marginally significant group \times congruency interaction was evident [$F_{(1,55)} = 3.82$, $p = 0.056$, $\eta^2 = 0.065$]. The Stroop effect (incongruent minus congruent) tended to be larger in the DD group ($M = 421.54$, $SD = 281.00$) than in the Control group ($M = 277.19$, $SD = 275.51$). In the Area task, a significant congruency was found [$F_{(1,55)} = 93.63$, $p < 0.001$, $\eta^2 = 0.630$], but no group effect [$F_{(1,55)} = 2.16$, $p = 0.147$, $\eta^2 = 0.038$]. No group \times congruency interaction was found [$F_{(1,55)} = 0.72$, $p = 0.397$, $\eta^2 = 0.013$]. In other words, in the Area task, there was a typical Stroop effect, and this pattern was similar for Controls ($M = 256.59$, $SD = 198.83$) and DD ($M = 306.27$, $SD = 236.10$). Similar to error rates analysis, the DD group tended to show larger Stroop effects (marginally significant) in the perimeter but not in the area task.

Proficiency Phase

While aiming to find out whether magnitude interference and group differences were evident after continuous exposure, results revealed a main effect of group [$F_{(1,55)} = 6.28$, $p = 0.015$, $\eta^2 = 0.103$]. DDs' RTs were significantly slower than Controls'. The main effect of congruency was also significant [$F_{(1,55)} = 241.8$, $p < 0.001$, $\eta^2 = 0.815$], such that participants responded more slowly to incongruent trials than to congruent trials. However, the main effect of task did not reach significance [$F_{(1,55)} = 0.16$, $p = 0.684$, $\eta^2 = 0.003$], such that participants responded similarly to Area and to Perimeter tasks.

Importantly and following the experimental hypothesis, the group \times congruency interaction was found to be significant [$F_{(1,55)} = 04.30$, $p = 0.043$, $\eta^2 = 0.073$]. The Stroop effect (incongruent minus congruent) was significantly larger [$F_{(1,55)} = 4.01$, $p = 0.043$, $\eta^2 = 0.068$] for the DD group ($M = 263.95$, $SD = 107.74$) than for the Control group ($M = 201.79$, $SD = 118.39$).

No triple interaction was found between group, congruency, and task in the proficiency phase [$F_{(1,55)} = 0.06$, $p = 0.794$, $\eta^2 = 0.001$]. However, following scientific interest, the double interaction between congruency and group was analyzed separately in each task. In the Perimeter task, significant congruency [$F_{(1,55)} = 120.68$, $p < 0.001$, $\eta^2 = 0.687$] and a marginally significant group effect [$F_{(1,55)} = 3.74$, $p = 0.058$,

$\eta^2 = 0.064$] were found. However, no group \times congruency interaction was evident [$F_{(1,55)} = 2.66$, $p = 0.108$, $\eta^2 = 0.046$] and both groups showed typical Stroop effects (DD: $M = 263.93$, $SD = 150.83$; Control: $M = 195.60$, $SD = 165.02$). Similarly, in the Area task, significant congruency [$F_{(1,55)} = 192.45$, $p < 0.001$, $\eta^2 = 0.778$] and group effects [$F_{(1,55)} = 8.25$, $p = 0.006$, $\eta^2 = 0.131$] were found. However, no group \times congruency interaction was found [$F_{(1,55)} = 2.7$, $p = 0.106$, $\eta^2 = 0.047$] and both groups showed typical Stroop effects (DD: $M = 263.97$, $SD = 144.12$; Control: $M = 207.98$, $SD = 107.81$). Accordingly, after continuous exposure to stimuli, both groups showed typical Stroop effects in both tasks.

Continuous Exposure Effect (i.e., The Statistical Differences between Stroop Effects in the Practice vs. Task Phases)

We analyzed the effect of exposure separately for each group in each task. This was done in order to figure out whether stimuli rehearsal led to changes in automaticity of one or both magnitudes among different participants. In the Perimeter task (in which Area is irrelevant to the task), a significant difference was found in the Stroop effects (congruent minus incongruent) between practice and proficiency phases for the DD group [$t_{(29)} = 3.64$, $p = 0.001$, $d = 1.351$]. Accordingly, high initial Stroop effects ($M = 421.54$, $SD = 281.006$) declined after practice ($M = 263.93$, $SD = 150.83$). No significant difference was evident between Stroop effects in the practice and proficiency phases for the Control group [$t_{(26)} = 1.62$, $p = 0.115$, $d = 0.635$; practice phase: $M = 277.19$, $SD = 275.51$; proficiency phase: $M = 195.60$, $SD = 165.02$]. No difference in Stroop effects was evident in the Area task (in which Perimeter was irrelevant to the task), neither for the DD group [$t_{(29)} = 1.17$, $p = 0.248$, $d = 0.434$; Practice phase: $M = 306.27$, $SD = 236.10$; Proficiency phase: $M = 263.97$, $SD = 144.12$], nor for the Control group [$t_{(26)} = 1.31$, $p = 0.202$, $d = 0.513$; practice phase: $M = 256.59$, $SD = 198.83$; proficiency phase: $M = 207.98$, $SD = 107.87$].

DISCUSSION

The present study investigated the automaticity of area and perimeter processing at different exposure levels in adults with deficient and intact numerical abilities.

To the best of our knowledge, this is the first study to show that individuals with DD process area and perimeter

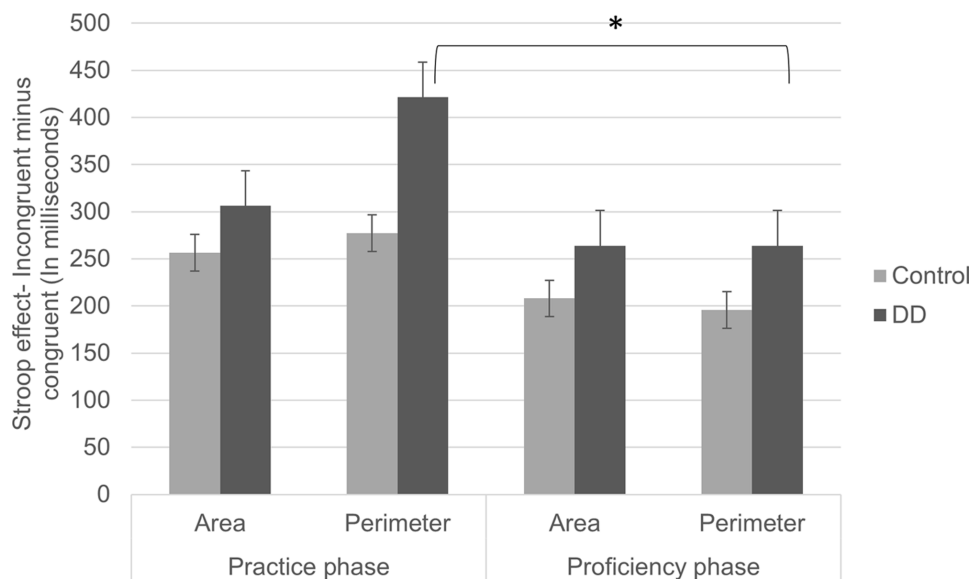


FIGURE 4 | Stroop effects (incongruent minus congruent trials) on RTs of practice and proficiency phases in both groups (DD and Control). Error bars depict ± 1 SE of the mean. Significance values: * $p < 0.05$.

information differently and in a less automatic manner than peers with intact numerical competence. After initial exposure to stimuli, area processing was more automatic than perimeter processing for both groups, as represented by task \times congruency interaction, evident in error rate analysis. Specifically, significant Stroop effects (slower responses to incongruent vs. congruent trials) were evident in perimeter and not in area tasks. However, this pattern was more prominent among the DD group, appearing in both error rates and speed analyses, implying a magnitude processing deficit. Together with previous evidence (Babai et al., 2006), the current findings show that area interferes with perimeter processing but not vice versa. This pattern suggests that while area processing may be innate, perimeter processing is acquired. After continuous exposure, the difference between area and perimeter was no longer evident. Furthermore, significant Stroop effects (in speed analysis) show that both magnitudes were automatically processed and interfered with each other to a similar degree in both groups. However, we found that domain general learning and inhibition deficits are also involved in DD. Specifically, we found firm group \times congruency interactions in the speed dimension. Namely, the DD group showed larger Stroop effects, which were evident across all exposure levels. Overall, the findings imply that deficient performance of participants with DD may not be restricted to numerical processing.

Automatic Processing of Area and Perimeter

In the current work, area and perimeter were both found to be automatically processed, as shown by the existence of significant Stroop effects (slower responses to incongruent vs. congruent

trials), evident from speed analysis in both area and perimeter tasks. In other words, the irrelevant aspect (area or perimeter) was involuntarily processed. As in the case of the numerical Stroop (Henik and Tzelgov, 1982), bidirectional effects were evident between the two magnitudes. The fact that both aspects are automatically processed is compatible with recent research highlighting the existence of a general magnitude processing mechanism (Gebuis and Reynvoet, 2012b; Newcombe, 2014; Leibovich et al., 2016).

Our findings are partially consistent with previous studies (Babai et al., 2006) showing that the area aspect interferes with perimeter judging but not vice versa. In the current study, perimeter processing seems to be automatic as well and to interfere with area processing after continuous exposure to the stimuli. However, some of our findings support the assumption that the area aspect is more perceptually salient. In the analysis of error rates in the practice phase we found Stroop effects (when congruent trials were more accurate than incongruent trials) in the perimeter but not in the area task. Area saliency probably relates to the fact that it occupies a larger space (in square meters) (Abbott, 1976). Moreover, magnitude saliency can differ due to task demands (Spelke et al., 2010). From this perspective, the processing of surface layouts (perimeter), which is crucial for navigation tasks (Cheng and Newcombe, 2005), is probably less vital than the figures' area in order to decide which small-scale 2D figure (Lee and Spelke, 2011) is larger, as required in the current task.

Since area saliency disappeared in the proficiency phase, we can assume that exposure levels also account for magnitude saliency. Accordingly, after practicing, participants were able to successfully inhibit irrelevant salient area magnitude information, and no difference was found between area and perimeter Stroop effects. Based on previous study that found

that as few as eight rehearsals trials, contributed to changes in brain functions (Ischebeck et al., 2007), we may argue, that in the current study which included larger number of stimuli, a meaningful learning indeed occurred. The findings are consistent with intervention studies indicating that attracting attention to the perimeter aspect (either by adding warnings or by presenting a polygon's perimeter in discrete units) improved participants' accuracy rates in similar geometric tasks (Babai et al., 2015, 2016). One should notice that the current research did not include direct comparison between numerical and continuous magnitudes. Accordingly, no conclusions regarding the interactions between area, perimeter, and numeracy in numerical tasks, such as dot arrays judgment tasks (as in Gebuis and Reynvoet, 2012a,b), can be made.

DD and Automatic Processing of Magnitudes

When faced with unfamiliar stimuli in the practice phase, a triple group \times task \times congruency interaction was evident in error rates analysis. Participants with DD showed larger Stroop effects compared to the control group in the perimeter task but not in the area task (as in Babai et al., 2006). This pattern was also marginally significant for speed analysis. These findings indicate that individuals with deficient numerical processing also struggle with magnitude, non-numerical processing (as suggested by Leibovich et al., 2016). Moreover, group differences were significant when trying to ignore the irrelevant but salient area magnitude (Babai et al., 2006; Stavy et al., 2006; Stavy and Babai, 2008, 2010). Accordingly, the findings suggest that individuals with DD are deficient not only in the processing of numerical stimuli (e.g., Kaufmann and von Aster, 2012), but also in the processing of continuous magnitudes (in this specific case, perimeter).

Nevertheless, with practice and growing proficiency in the tasks' demands, no triple interactions were evident. The difference between perimeter and area tasks disappeared and both groups showed typical Stroop effects for both magnitudes. In other words, both magnitudes were processed automatically. Accordingly, group differences in magnitude comparison tasks may vary due to task familiarity or proficiency levels. Developmental studies regarding the numeric Stroop task indicate that one magnitude (physical size) seems to be more salient and to interfere with the other (symbolic numbers) among first graders who have no formal numerical education. With educational progress and repeated experience with numbers, both magnitudes are automatically processed and bidirectional effects are evident (Girelli et al., 2000; Rubinsten et al., 2002). Here we show a similar pattern, as task practice resulted in bidirectional effects for both groups and both magnitudes interfered with each other's judgment.

Further analysis revealed that the reduction of Stroop effects in the perimeter task was more prominent for the DD group. Accordingly, participants with DD showed automatic processing of the perimeter aspect only after continuous exposure to stimuli. Namely, they had to intentionally learn to process the perimeter aspect. The findings emphasize the need to summon intensive

exposure to magnitude, as well as to numerical information, in order to enhance compensation of DDs' core deficits (e.g., Kaufmann et al., 2003; Wilson et al., 2006).

Beyond the earlier discussion regarding magnitude saliency, a group \times congruency interaction was evident on the speed dimension across all exposure phases. Participants with DD systematically struggled to inhibit irrelevant magnitude information, manifested by larger Stroop effects, regardless of task type – area or perimeter. This pattern is compatible with inhibition deficits often associated with DD (Wang et al., 2012; Szucs et al., 2013a; Bugden and Ansari, 2015). As suggested by Aubin et al. (2016), individuals with DD might fail to consolidate magnitude knowledge and to produce the shift from the “slow” computing neural network to the “fast” retrieval network. Hence, they must invest more cognitive efforts in order to ignore irrelevant magnitude information. According to the magnitude mechanism hypothesis, a separate and more precise representation of each magnitude dimension occurs across development and becomes stronger with formal education (Newcombe, 2014). In line with this theory, people with DD struggle to effectively process different magnitude aspects of the stimulus in order to extract the proper magnitude on one hand and ignore the irrelevant magnitude on the other (Gebuis and Gevers, 2011). Hence, individuals with DD may show a developmental gap (Rubinsten and Henik, 2005), demonstrated by a difficulty in differentiating between magnitudes. Based on the current findings, we cannot specify whether individuals with DD have a deficient, domain-general inhibition mechanism (Soltész et al., 2007), or whether these inhibition deficits are exclusive to magnitude processing (De Visscher and Noël, 2014). It is also plausible that multiple neuro-cognitive components, domain-specific and domain-general, account for DD (Rubinsten and Henik, 2009; McCaskey et al., 2017). For instance, some researchers argue that DD relates to visual-spatial deficits (Ashkenazi et al., 2013; Szucs et al., 2013a; Bugden and Ansari, 2015). The fact that the DD group performed worse than the control group on reading tasks may strengthen this argument, as the reading process is known to involve visual-spatial processing (Cohen et al., 2000; Facoetti et al., 2000).

Since assessment tools and inclusion criteria vary between studies concerning numerical difficulties (Kaufmann et al., 2013), the debate on which domain-specific and domain-general deficits underlie numerical difficulties has not yet been fully addressed. There is also a need to develop a massive body of research involving developmental research and using other continuous magnitude assessment tasks. However, it is worth noting that, based on developmental research, Leibovich and Henik (2014), Leibovich et al. (2016) suggest a developmental model of the magnitude processing system: at the first step infants process mostly continuous magnitudes (for example, Clearfield and Mix, 1999). With the developmental process, they learn about the correlation between continuous and discrete magnitudes (numbers) and are subsequently able to process both types of magnitude separately (Leibovich and Henik, 2014). Development of cognitive control, including inhibition mechanisms (Davidson et al., 2006), is crucial in order to enable differentiation between

incongruent magnitudes (Leibovich et al., 2016). The proposed model has not yet been systematically validated, and the study of possible developmental gaps that might shed light on how numerical difficulties arise and develop, has not been established. However, the current study implies that DD deficits indeed relate to a failure to differentiate between magnitudes and to inhibit the irrelevant magnitude.

Other Visual Properties

The uniqueness of the current Stroop-like task enabled us to look at the automaticity of framed vs. filled figures separately. Results show that participants showed larger Stroop effects in framed vs. filled trials, regardless of trial type (area or perimeter) and group (DD or control). These findings are not consistent with previous ones suggesting that framed figures emphasize the perimeter aspect and thus help ignore irrelevant area information (Stavy and Babai, 2008). Magnitudes tend to become confounded in a natural environment (Gebuis and Reynvoet, 2012a,b). Accordingly, it is possible that filled figures, occupying as they do larger magnitudes than framed figures, increased area and perimeter's natural confound.

Additional visual variables were examined in order to eliminate possible alternative explanations of the results. While no full control of other visual features is possible (Gebuis and Reynvoet, 2012a,b), changes in these variables appear to be good indicators that participants' performance was affected by the manipulation of experimental variables (i.e., area and perimeter). No group differences were found in any of the other visual features, including protrusion and vertical or horizontal direction. Hence, we can argue that group differences arise from magnitude processing rather than from interference of irrelevant visual features.

Limitations

One limitation of the current study is that all participants were well-educated students and hence might not be representative of the entire DD demographic. Furthermore, the relationship between magnitude and numerical processing was not directly assessed in the current research, a fact that reduces our ability to generate firm conclusions about the role of magnitude processing in numerical processing.

CONCLUSION AND IMPLICATIONS

The current study indicates that both area and perimeter magnitudes are automatically processed and that participants

with DD find it harder to ignore irrelevant magnitude information, especially salient area information. We assume that our findings derive from an inhibition deficit related to magnitude processing (Wang et al., 2012; De Visscher and Noël, 2014). The findings are also compatible with recent theories regarding the general magnitude processing mechanism (Leibovich and Henik, 2014; Newcombe, 2014; Lourenco and Bonny, 2016; Leibovich et al., 2016). On the other hand, we show that continuous exposure to stimuli was effective and resulted in similar patterns, namely typical Stroop effects for both magnitudes, for both groups. This fact is important for the planning of future intervention programs emphasizing the vitality of massive exposure to magnitude related stimuli in order to overcome the core deficits related to DD.

One main conclusion that can be deduced from the above is the need to develop other tasks assessing non-numerical magnitude processing, such as the task presented in the current work. Investigating multiple magnitude processing is crucial for both research and educational fields. On the one hand, there is a need for better knowledge about how multiple magnitude processing develops and occurs on the neuronal and behavioral levels in order to develop adaptive behavior. Traditionally, research methods have aimed to assess numerical processing exclusively and to block interfering magnitude information. On the other hand, the current and recent works (Gebuis and Reynvoet, 2012b; Szűcs et al., 2013b; Newcombe, 2014) stress that numerical and magnitude confounding should receive more attention in order to understand intact and impaired numerical processing. Furthermore, since non-numerical magnitude processing is crucial for higher education in science and mathematics (Wai et al., 2009; Marghetis et al., 2016), it is necessary to relate to non-numerical magnitude processing in educational research and in math curricula.

AUTHOR CONTRIBUTIONS

HE-L and OR contributed significantly to research design, data interpretation, and manuscript drafting. HE-L collected the data.

ACKNOWLEDGMENTS

We wish to acknowledge the help of Dr. Reuven Babai and Prof. Ruth Stavy in developing the Stroop-like tasks presented in the current study and adjusting the statistical model.

REFERENCES

- Abbott, P. (1976). *Geometry*. London: Hodder and Stoughton.
- Ashkenazi, S., and Henik, A. (2010). Attentional networks in developmental dyscalculia. *Behav. Brain Funct.* 6:2. doi: 10.1186/1744-9081-6-2
- Ashkenazi, S., Rosenberg-Lee, M., Metcalfe, A. W. S., Swigart, A. G., and Menon, V. (2013). Neuropsychologia visuo – spatial working memory is an important source of domain-general vulnerability in the development of arithmetic cognition. *Neuropsychologia* 51, 2305–2317. doi: 10.1016/j.neuropsychologia.2013.06.031
- Ashkenazi, S., Rubinsten, O., and Henik, A. (2009). Attention, automaticity, and developmental dyscalculia. *Neuropsychology* 23, 535–540. doi: 10.1037/a0015347
- Aubin, S., Voelker, A. R., and Eliasmith, C. (2016). Improving with practice: a neural model of mathematical development. *Top. Cogn. Sci.* 9, 6–20. doi: 10.1111/tops.12242
- Babai, R., Levyadun, T., Stavy, R., and Tirosh, D. (2006). Intuitive rules in science and mathematics: a reaction time study. *Int. J. Math. Educ. Sci. Technol.* 37, 913–924. doi: 10.1080/002073906000794958

- Babai, R., Nattiv, L., and Stavy, R. (2016). Comparison of perimeters: improving students' performance by increasing the salience of the relevant variable. *ZDM* 48, 367–378. doi: 10.1007/s11858-016-0766-z
- Babai, R., Shalev, E., and Stavy, R. (2015). A warning intervention improves students' ability to overcome intuitive interference. *ZDM* 47, 735–745. doi: 10.1007/s11858-015-0670-y
- Bugden, S., and Ansari, D. (2015). Probing the nature of deficits in the 'Approximate Number System' in children with persistent developmental dyscalculia. *Dev. Sci.* 19, 817–833. doi: 10.1111/desc.12324
- Butterworth, B. (2005). "Developmental dyscalculia," in *Handbook of Mathematical Cognition*, ed. J. I. D. Campbell (New York, NY: Psychology Press), 455–467.
- Cheng, K., and Newcombe, N. (2005). Is there a geometric module for spatial orientation? Squaring theory and evidence. *Psychon. Bull. Rev.* 12, 1–23. doi: 10.3758/BF03196346
- Clearfield, M. W., and Mix, K. S. (1999). Number versus contour length in infants' discrimination of small visual sets. *Psychol. Sci.* 10, 408–411. doi: 10.1111/1467-9280.00177
- Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M. A., et al. (2000). The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain* 123, 291–307. doi: 10.1093/brain/123.2.291
- Dakin, S. C., Tibber, M. S., Greenwood, J. A., Kingdom, F. A. A., and Morgan, M. J. (2011). A common visual metric for approximate number and density. *Proc. Natl. Acad. Sci. U.S.A.* 108, 19552–19557. doi: 10.1073/pnas.1113195108
- Davidson, M. C., Amso, D., Anderson, L. C., and Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia* 44, 2037–2078. doi: 10.1016/j.neuropsychologia.2006.02.006
- De Visscher, A., and Noël, M.-P. (2014). The detrimental effect of interference in multiplication facts storing: typical development and individual differences. *J. Exp. Psychol. Gen.* 143, 2380–2400. doi: 10.1037/xge0000029
- Dehaene, S., and Brannon, E. M. (2010). Space, time, and number: a Kantian research program. *Trends Cogn. Sci.* 14, 517–519. doi: 10.1016/j.tics.2010.09.009
- Dehaene, S., Piazza, M., Pinel, P., and Cohen, L. (2003). Three parietal circuits for number processing. *Cogn. Neuropsychol.* 20, 487–506. doi: 10.1080/02643290244000239
- Derdikman, D., and Moser, E. I. (2010). A manifold of spatial maps in the brain. *Trends Cogn. Sci.* 14, 561–569. doi: 10.1016/j.tics.2010.09.004
- Desoete, A., Stock, P., Schepens, A., Baeyens, D., and Roeyers, H. (2009). Classification, seriation, and counting in grades 1, 2, and 3 as two-year longitudinal predictors for low achieving in numerical facility and arithmetical achievement? *J. Psychoeduc. Assess.* 27, 252–264. doi: 10.1177/0734282908330588
- Facoetti, A., Paganoni, P., Turatto, M., Marzola, V., and Mascetti, G. G. (2000). Visual-spatial attention in developmental dyslexia. *Cortex* 36, 109–123. doi: 10.1016/S0010-9452(08)70840-2
- Feigenson, L., Carey, S., and Spelke, E. S. (2002). Infants' discrimination of number vs. continuous extent. *Cogn. Psychol.* 44, 33–66. doi: 10.1006/cogp.2001.0760
- Geary, D. C. (1993). Mathematical disabilities: cognitive, neuropsychological, and genetic components. *Psychol. Bull.* 114, 345–362. doi: 10.1037/0033-2909.114.2.345
- Gebuis, T., and Gevers, W. (2011). Numerosities and space; indeed a cognitive illusion! A reply to de Hevia and Spelke (2009). *Cognition* 121, 248–52; discussion 253–255. doi: 10.1016/j.cognition.2010.09.008
- Gebuis, T., and Reynvoet, B. (2012a). Continuous visual properties explain neural responses to nonsymbolic number. *Psychophysiology* 49, 1481–1491. doi: 10.1111/j.1469-8986.2012.01461.x
- Gebuis, T., and Reynvoet, B. (2012b). The neural mechanisms underlying passive and active processing of numerosity. *NeuroImage* 70:301–307. doi: 10.1016/j.neuroimage.2012.12.048
- Girelli, L., Lucangeli, D., and Butterworth, B. (2000). The development of automaticity in accessing number magnitude. *J. Exp. Child Psychol.* 76, 104–122. doi: 10.1006/jecp.2000.2564
- Henik, A., and Tzelgov, J. (1982). Is three greater than five: the relation between physical and semantic size in comparison tasks. *Mem. Cogn.* 10, 389–395. doi: 10.3758/BF03202431
- Hermer, L., and Spelke, E. (1994). A geometric process for spatial representation in young children. *Nature* 370, 57–59. doi: 10.1016/j.jecp.2009.05.007
- Hermer-Vazquez, L., Spelke, E., and Katsnelson, A. (1999). Source of flexibility in human cognition: dual task studies of space and language. *Cognit. Psychol.* 39, 3–36. doi: 10.1006/cogp.1998.0713
- Huang, Y., and Spelke, E. S. (2015). Core knowledge and the emergence of symbols: the case of maps. *J. Cogn. Dev.* 6, 81–96. doi: 10.1080/15248372.2013.784975
- Huber, S., Sury, D., Moeller, K., Rubinsten, O., and Nuerk, H. C. (2015). A general number-to-space mapping deficit in developmental dyscalculia. *Res. Dev. Disabil.* 43, 32–42. doi: 10.1016/j.ridd.2015.06.003
- Ischebeck, A., Zamarian, L., Egger, K., Schocke, M., and Delazer, M. (2007). Imaging early practice effects in arithmetic. *Neuroimage* 36, 993–1003. doi: 10.1016/j.neuroimage.2007.03.051
- Kaufmann, L., Handl, P., and Thöny, B. (2003). Evaluation of a numeracy intervention program focusing on basic numerical knowledge and conceptual knowledge: a pilot study. *J. Learn. Disabil.* 36, 564–573. doi: 10.1177/00222194030360060701
- Kaufmann, L., Mazzocco, M. M., Dowker, A., von Aster, M., Göbel, S. M., Grabner, R. H., et al. (2013). Dyscalculia from a developmental and differential perspective. *Front. Psychol.* 4:516. doi: 10.3389/fpsyg.2013.00516
- Kaufmann, L., and von Aster, M. (2012). The diagnosis and management of dyscalculia. *Deutsch. Ärztebl. Int.* 109, 767–777.
- Lee, S. A., Sovrano, V. A., and Spelke, E. S. (2012). Navigation as a source of geometric knowledge: young children's use of length, angle, distance, and direction in a reorientation task. *Cognition*, 123, 144–161. doi: 10.1016/j.cognition.2011.12.015
- Lee, S. A., and Spelke, E. S. (2011). Young children reorient by computing layout geometry, not by matching images of the environment. *Psychon. Bull. Rev.* 18, 192–198. doi: 10.3758/s13423-010-0035-z
- Leibovich, T., and Henik, A. (2013). Magnitude processing in non-symbolic stimuli. *Front. Psychol.* 4:375. doi: 10.3389/fpsyg.2013.00375
- Leibovich, T., and Henik, A. (2014). Comparing performance in discrete and continuous comparison tasks. *Q. J. Exp. Psychol.* 67, 899–917. doi: 10.1080/17470218.2013.837940
- Leibovich, T., Katzin, N., Harel, M., and Henik, A. (2016). From 'sense of number' to 'sense of magnitude' – The role of continuous magnitudes in numerical cognition. *Behav. Brain Sci.* doi: 10.1017/S0140525X16000960 [Epub ahead of print].
- Lourenco, S. F., and Bonny, J. W. (2016). Representations of numerical and non-numerical magnitude both contribute to mathematical competence in children. *Dev. Sci.* 20:e12418. doi: 10.1111/desc.12418
- Lourenco, S. F., and Longo, M. R. (2010). General magnitude representation in human infants. *Psychol. Sci.* 21, 873–881. doi: 10.1177/0956797610370158
- Marghetis, T., Landy, D., and Goldstone, R. L. (2016). Mastering algebra retrains the visual system to perceive hierarchical structure in equations. *Cogn. Res.* 1:25. doi: 10.1186/s41235-016-0020-9
- Mazzocco, M. M., Devlin, K. T., and McKenney, S. J. (2008). Is it a fact? Timed arithmetic performance of children with mathematical learning disabilities (MLD) varies as a function of how MLD is defined. *Dev. Neuropsychol.* 33, 318–344. doi: 10.1080/87565640801982403
- McCaskey, U., von Aster, M., O'Gorman Tuura, R., and Kucian, K. (2017). Adolescents with developmental dyscalculia do not have a generalized magnitude deficit-processing of discrete and continuous magnitudes. *Front. Hum. Neurosci.* 11:102. doi: 10.3389/fnhum.2017.00102
- Menon, V., Rivera, S. M., White, C. D., Glover, G. H., and Reiss, A. L. (2000). Dissociating prefrontal and parietal cortex activation during arithmetic processing. *Neuroimage* 12, 357–365. doi: 10.1006/nimg.2000.0613
- Mussolin, C., De Volder, A., Grandin, C., Schlögel, X., Nassogne, M.-C., and Noël, M.-P. (2010a). Neural correlates of symbolic number processing in developmental dyscalculia. *J. Cogn. Neurosci.* 22, 860–874. doi: 10.1162/jocn.2009.21237
- Mussolin, C., Mejias, S., and Noel, M. P. (2010b). Symbolic and nonsymbolic number comparison in children with and without dyscalculia. *Cognition* 115, 10–25. doi: 10.1016/j.cognition.2009.10.006
- Newcombe, N. S. (2014). The origins and development of magnitude estimation. *Ecol. Psychol.* 26, 147–157. doi: 10.1080/10407413.2014.875333
- Noël, M. P., Rousselle, L., and De Visscher, A. (2016). Both specific and general cognitive factors account for dyscalculia. *Spec. Needs Math. Educ.* 8, 35–52. doi: 10.1177/0022219414547221

- Odic, D., Libertus, M. E., Feigenson, L., and Halberda, J. (2012). Developmental change in the acuity of approximate number and area representations. *Dev. Psychol.* 49, 1103–1112. doi: 10.1037/a0029472
- Pinel, P., Piazza, M., Le Bihan, D., and Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron* 41, 983–993. doi: 10.1016/S0896-6273(04)00107-2
- Price, G. R., Holloway, I., Rasanen, P., Vesterinen, M., and Ansari, D. (2007). Impaired parietal magnitude processing in developmental dyscalculia. *Curr. Biol.* 17, 1042–1043. doi: 10.1016/j.cub.2007.10.013
- Raven, J. C., Court, J. H., and Raven, J. (1995). *Raven, Matrices Progresivas (Escalas: CPM, SPM, APM)*. Madrid: TEA Ediciones, S.A.
- Reigosa-Crespo, V., Valdés-Sosa, M., Butterworth, B., Estévez, N., Rodríguez, M., Santos, E., et al. (2012). Basic numerical capacities and prevalence of developmental dyscalculia: the Havana Survey. *Dev. Psychol.* 48, 123–135. doi: 10.1037/a0025356
- Ren, P., Nicholls, M. E. R., Ma, Y., and Chen, L. (2011). Size matters: non-numerical magnitude affects the spatial coding of response. *PLOS ONE* 6:e23553. doi: 10.1371/journal.pone.0023553
- Rotzer, S., Kucian, K., Martin, E., von Aster, M., Klaver, P., and Loenneker, T. (2008). Optimized voxel-based morphometry in children with developmental dyscalculia. *NeuroImage* 39, 417–422. doi: 10.1016/j.neuroimage.2007.08.045
- Rubinsten, O., and Henik, A. (2005). Automatic activation of internal magnitude: a study of developmental dyscalculia. *Neuropsychology* 19, 641–648. doi: 10.1037/0894-4105.19.5.641
- Rubinsten, O., and Henik, A. (2006). Double dissociation of functions in developmental dyslexia and dyscalculia. *J. Educ. Psychol.* 98, 854–867. doi: 10.1037/0022-0663.98.4.854
- Rubinsten, O., and Henik, A. (2009). Developmental dyscalculia: heterogeneity might not mean different mechanisms. *Trends Cogn. Sci.* 13, 92–99. doi: 10.1016/j.tics.2008.11.002
- Rubinsten, O., Henik, A., Berger, A., and Shahar-Shalev, S. (2002). The development of internal representations of magnitude and their association with Arabic numerals. *J. Exp. Child Psychol.* 81, 74–92. doi: 10.1006/jecp.2001.2645
- Rubinsten, O., and Sury, D. (2011). Processing ordinality and quantity: the case of developmental dyscalculia. *PLOS ONE* 6:e24079. doi: 10.1371/journal.pone.0024079
- Shalev, R. S., and Gross-Tsur, V. (2001). Developmental Dyscalculia. *Pediatr. Neurol.* 24, 337–342. doi: 10.1016/S0887-8994(00)00258-7
- Skagerlund, K., and Träff, U. (2014). Development of magnitude processing in children with developmental dyscalculia: space, time, and number. *Front. Psychol.* 5:675. doi: 10.3389/fpsyg.2014.00675
- Soltész, F., Szucs, D., Dékány, J., Márkus, A., and Csépe, V. (2007). A combined event-related potential and neuropsychological investigation of developmental dyscalculia. *Neurosci. Lett.* 417, 181–186. doi: 10.1016/j.neulet.2007.02.067
- Spelke, E., Lee, S. A., and Izard, V. (2010). Beyond core knowledge: natural geometry. *Cogn. Sci.* 34, 863–884. doi: 10.1111/j.1551-6709.2010.01110.x
- Stavy, R., and Babai, R. (2008). Complexity of shapes and quantitative reasoning in geometry. *Mind Brain Educ.* 2, 170–176. doi: 10.1111/j.1751-228X.2008.00051.x
- Stavy, R., and Babai, R. (2010). Overcoming intuitive interference in mathematics: insights from behavioral, brain imaging, and intervention studies. *Int. J. Math. Educ.* 42, 621–633. doi: 10.1007/s11858-010-0251-z
- Stavy, R., Goel V., Critchley H., and Dolan, R. (2006). Intuitive interference in quantitative reasoning. *Brain Res.* 1073–1074, 383–388. doi: 10.1016/j.brainres.2005.12.011
- Stoianov, I., and Zorzi, M. (2012). Emergence of a “visual number sense” in hierarchical generative models. *Nat. Neurosci.* 15, 194–196. doi: 10.1038/nn.2996
- Szucs, D., Devine, A., Soltesz, F., Nobes, A., and Gabriel, F. (2013a). Developmental dyscalculia is related to visuo-spatial memory and inhibition impairment. *Cortex* 49, 2674–2688. doi: 10.1016/j.cortex.2013.06.007
- Szűcs, D., Nobes, A., Devine, A., Gabriel, F. C., and Gebuis, T. (2013b). Visual stimulus parameters seriously compromise the measurement of approximate number system acuity and comparative effects between adults and children. *Front. Psychol.* 4:444. doi: 10.3389/fpsyg.2013.00444
- Szűcs, D., and Goswami, U. (2013). Developmental dyscalculia: fresh perspectives. *Trends Neurosci. Educ.* 2, 33–37. doi: 10.1016/j.tine.2013.06.004
- Tibber, M. S., Greenwood, J. A., and Dakin, S. C. (2012). Number and density discrimination rely on a common metric: Similar psychophysical effects of size, contrast, and divided attention. *J. Vis.* 12:8. doi: 10.1167/12.6.8
- Träff, U., Olsson, L., Östergren, R., and Skagerlund, K. (2016). Heterogeneity of developmental dyscalculia: cases with different deficit profiles. *Front. Psychol.* 7:2000. doi: 10.3389/fpsyg.2016.02000
- Tzelgov, J. (1997). “Automatic but conscious: that is how we act most of the time,” in *Advances in Social Cognition*, Vol. 10, ed. R. S. Wyer, Jr (Mahwah, NJ: Erlbaum), 217–230.
- Wai, J., Lubinski, D., and Benbow, C. P. (2009). Spatial ability for STEM domains: aligning over 50 years of cumulative psychological knowledge solidifies its importance. *J. Educ. Psychol.* 101, 817–835. doi: 10.1037/a0016127
- Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space, and quantity. *Trends Cogn. Sci.* 7, 483–488. doi: 10.1016/j.tics.2003.09.002
- Wang, L. C., Tasi, H. J., and Yang, H. M. (2012). Cognitive inhibition in students with and without dyslexia and dyscalculia. *Res. Dev. Disabil.* 33, 1453–1461. doi: 10.1016/j.ridd.2012.03.019
- Wilson, A. J., Revkin, S. K., Cohen, D., Cohen, L., and Dehaene, S. (2006). An open trial assessment of “The Number Race,” an adaptive computer game for remediation of dyscalculia. *Behav. Brain Funct.* 2:20.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Eidlin-Levy and Rubinsten. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Are Books Like Number Lines? Children Spontaneously Encode Spatial-Numeric Relationships in a Novel Spatial Estimation Task

Clarissa A. Thompson^{1†}, Bradley J. Morris^{2†} and Pooja G. Sidney¹

¹ Department of Psychological Sciences, Kent State University, Kent, OH, United States, ² Department of Educational Psychology, Kent State University, Kent, OH, United States

OPEN ACCESS

Edited by:

Hans-Christoph Nuerk,
Universität Tübingen, Germany

Reviewed by:

Koen Luwel,
KU Leuven, Belgium
Jo Van Herwegen,
Kingston University, United Kingdom

*Correspondence:

Clarissa A. Thompson
cthomp77@kent.edu

[†] These authors have contributed
equally to this work and shared first
authorship.

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 06 October 2017

Accepted: 11 December 2017

Published: 21 December 2017

Citation:

Thompson CA, Morris BJ and
Sidney PG (2017) Are Books Like
Number Lines? Children
Spontaneously Encode
Spatial-Numeric Relationships
in a Novel Spatial Estimation Task.
Front. Psychol. 8:2242.
doi: 10.3389/fpsyg.2017.02242

Do children spontaneously represent spatial-numeric features of a task, even when it does not include printed numbers (Mix et al., 2016)? Sixty first grade students completed a novel spatial estimation task by seeking and finding pages in a 100-page book without printed page numbers. Children were shown pages 1 through 6 and 100, and then were asked, “Can you find page X?” Children’s precision of estimates on the page finder task and a 0-100 number line estimation task was calculated with the Percent Absolute Error (PAE) formula (Siegler and Booth, 2004), in which lower PAE indicated more precise estimates. Children’s numerical knowledge was further assessed with: (1) numeral identification (e.g., What number is this: 57?), (2) magnitude comparison (e.g., Which is larger: 54 or 57?), and (3) counting on (e.g., Start counting from 84 and count up 5 more). Children’s accuracy on these tasks was correlated with their number line PAE. Children’s number line estimation PAE predicted their page finder PAE, even after controlling for age and accuracy on the other numerical tasks. Children’s estimates on the page finder and number line tasks appear to tap a general magnitude representation. However, the page finder task did not correlate with numeral identification and counting-on performance, likely because these tasks do not measure children’s magnitude knowledge. Our results suggest that the novel page finder task is a useful measure of children’s magnitude knowledge, and that books have similar spatial-numeric affordances as number lines and numeric board games.

Keywords: spatial-numeric association, numerical representation, magnitude knowledge, number line estimation, numeracy, literacy

INTRODUCTION

Do children spontaneously represent spatial-numeric features of a task, even when it does not include printed numbers (Mix et al., 2016)? Previous research has provided evidence of spatial-numeric associations early in development suggesting that space and number share a common representational format (McCrink and Opfer, 2014; Patro et al., 2014). We investigated the possibility that books have spatial-numeric affordances like number lines and board games. Specifically, all three share left-to-right orientations, promote equal spacing between values, and provide an explicit means for mapping numbers to relative magnitudes (see **Table 1**). The

overarching goal of the present experiment was to compare children's performance on tasks known to tap numerical knowledge to a novel measure, the page finder task, which asked children to estimate the location of a page within a book without labeled page numbers and is hypothesized to measure magnitude estimation. If book affordances are related to the affordances of other measures, such as number lines, then the results of the page finder task should be highly related to other measures that tap children's numerical magnitude understanding.

Numbers and Space

Number sense refers to representing and processing numbers and includes several underlying processes, such as the ability to subitize a small number of items exactly, count, and compare approximate values (Dehaene, 2011; Friso-van den Bos et al., 2014). Children's number sense becomes formalized as they map number words onto a mental number line via cultural tools (e.g., number lines; Thompson and Opfer, 2010). As children get older and gain more experience with numbers, they increasingly differentiate the underlying spatial-numeric representations into more precise number concepts (e.g., 75 is bigger than 35; Siegler, 2016), and this precision is predictive of concurrent and future performance on standardized mathematics achievement tests (Siegler et al., 2011; Starr et al., 2013; Fazio et al., 2014; Siegler and Thompson, 2014).

Numbers are represented both as approximate magnitudes and as exact categories such as "five" (Dehaene, 2011). Comparisons of approximate magnitudes are faster and more accurate as the ratio of difference between numbers increases (e.g., the numerical distance effect), and this provides evidence for spatial-numeric associations (Dehaene, 2011). According to the numerical distance effect, participants are faster and more accurate when deciding that 4 is larger than 1 than when deciding that 3 is larger than 2 because the mental representations for 4 and 1 overlap to a lesser degree than do the mental representations for 3 and 2. Thus, 4 and 1 are more distant and discriminable from one another than are 3 and 2.

Approximate number magnitudes are represented in a left-to-right ascending order along a mental number line in which small numbers are oriented on the left and large numbers are oriented on the right (Siegler and Opfer, 2003; Siegler, 2016). Evidence for spatial-numeric associations in children (van Galen and Reitsma, 2008), adults (Fias, 2001), and even chimpanzees (Adachi, 2014), comes from the investigation of the SNARC effect (Spatial Numerical Association of Response Codes) in which response rates are faster for relatively small numbers (0–4) when responses are made with the left hand and faster for large numbers (5–9) when responses are made with the right hand (Dehaene et al., 1993; Wood et al., 2008). The SNARC effect demonstrates a response bias consistent with a mental number line in which numbers increase in magnitude from left-to-right in cultures with left-to-right orthographies (Dehaene, 2011).

As further evidence of the spatial-numeric association in children, even young preschoolers show an advantage on numerical tasks that have an orientation that is consistent with the left-to-right directionality of writing in their cultures. For

instance, United States children played a spatial search match-to-sample game in which they were shown two boxes with seven compartments each. The compartments in the sample and matching box were verbally labeled in an increasing numeric order from left-to-right or right-to-left. In the game, children were shown an object hidden in one of the compartments in the sample box, and they were asked to find another object that was hidden in the same numbered compartment in the matching box. Children were faster and more accurate at finding the hidden object in the matching box if both boxes were verbally numbered from left-to-right as compared to right-to-left (Opfer et al., 2010). Further, those children who spontaneously counted an array of ten chips from left-to-right, added one chip to the right side of a row of three chips, and took away one chip from the right side of a row of four chips were more likely to accurately give a researcher a specified number of chips in the typical Give-N task (e.g., Can you give me 8 chips?) as compared to those children who did not display this spatial-numeric association (Opfer et al., 2010).

Spatial-Numeric Features of the Number Line and Cues That Co-vary with Number

Given the spatial-numeric nature of children's numerical representations (i.e., the mental number line), the number line estimation task has emerged as a robust (e.g., Laski and Siegler, 2007; Booth and Siegler, 2008; Opfer and Thompson, 2008; Thompson and Opfer, 2008, 2010, 2016) and predictive (e.g., Booth and Siegler, 2006; Siegler et al., 2011, 2012; Siegler and Thompson, 2014) measure of children's underlying numerical representations. In the number line estimation task, participants are shown a left endpoint labeled with 0 and a right endpoint labeled with a much larger number, such as 100. Participants' job is to estimate the location of a third number on the line by making a vertical hatch mark. Initially, numerical representations, as measured by the number line estimation task, are characterized by even (i.e., linear) spacing across smaller numeric ranges and compression across larger numeric ranges (see Siegler et al., 2009 for a review). For instance, second graders make accurate, linear estimates in the 0–100 range and less precise estimates in the 0–1,000 range (Siegler and Opfer, 2003). These children are not only more accurate in their small-scale estimates, they are also more *confident* in their small-scale as opposed to large-scale estimates (Wall et al., 2016). As children gain experience or receive corrective feedback on their estimates, they show linear spacing across increasingly larger numeric ranges (Opfer and Siegler, 2007; Opfer and Thompson, 2008; Thompson and Opfer, 2008, 2016), however, even adults continue to struggle to produce linear estimates in some very large numeric ranges. That is, only about half of adults make accurate, linearly spaced estimates in the 0 – billion numeric range (Landy et al., 2013). It should be noted that there has been a recent debate about the shape of children's numerical representations, and proponents of the proportion judgment account (e.g., Barth and Paladino, 2011; Slusser et al., 2013) suggest that a cyclical power function fits children's estimates better than

TABLE 1 | Comparison of magnitude affordances across materials.

Affordance	Number line	Linear board game	Books
Orientation: Magnitudes increase from left-to-right	Left side is zero value and right side is maximum value (e.g., 100)	Starting space (on left side) is one and final space is maximum value (e.g., 10 or 100)	As pages are flipped, pages with smaller numbers are placed on the left and pages with larger numbers remain on the right
Linearity/Movement: Equal distance between moves/Individual moves represent the same distance	Each hatch mark represents equal value/moving from five to six on the number line is equivalent to moving from 55 to 56	Each space on game board represents one value/Each space is an equivalent move	Each page represents two values (front and back)/Each page turn is an equivalent move
Spatial/Temporal: Increasing physical space between locations (or time to reach location) indicates larger magnitudes	When starting from 0, finding and marking larger numbers (e.g., 73) takes longer than smaller numbers (e.g., 11), and there is a larger physical distance between 0 and the larger number (e.g., 11 units between 0 and 11 and 73 between 0 and 73).	Moving to spaces farther from the initial space takes longer amounts of time than moving to spaces closer to the initial space	From first page, finding a page with a larger number (e.g., 73) takes a longer time than finding a page with a smaller number (e.g., 11)

a logarithmic or a linear function. However, proponents of the logarithmic-to-linear shift account (Opfer et al., 2011, 2016) suggest that providing children with feedback about the number located at the midpoint of a 0–1,000 number line anchors their estimates to 500, thus making the fit of the cyclical power function to children's number line estimates an artifact of the experimental methodology used. In the current paper, however, it is not our goal to make claims about children's conceptual change in number line estimation tasks (e.g., best fitting function that characterizes children's underlying numerical representation).

The number line task has both spatial and numeric components. There are numerically labeled end-points on the number line as well as a to-be-estimated number that appears above the number line. To estimate the magnitudes appropriately, the child needs to map the to-be-estimated number to the correct spatial location (i.e., distance from the left and right end point). Sidney et al. (2017, see Figure 1 from their paper) suggest that, at a minimum, children must employ cross-format proportional reasoning to make accurate, linear estimates of where given numbers are located on number lines, for example, in a typical number-to-position task (Siegler and Opfer, 2003). In this task, children are shown a line segment with symbolic anchors of 0 and 100 at the endpoints, and children's job is to find where along the line the to-be-estimated number is located. For instance, to accurately place 78 on a 0–100 number line, a child must estimate the length of a line segment that is 78% of the distance of the 100-unit line. To do so, a child must consider the ratio of the numerical magnitudes of 78 and 100 and match that ratio to the spatial magnitude of the 0–100 number line to estimate the spatial magnitude of a 0–78 line segment (see Barth and Paladino, 2011). The number line estimation task is a prime example of how space and number are naturally integrated. To accurately complete the number line estimation task, participants must map an internal numerical magnitude representation to an external physical location on the line. Children who have a more precise mapping between their internal numerical magnitude

representation and external spatial extent make more accurate number line estimates.

Improving Number Sense

Improving children's estimates on the number line task appears to improve the precision of children's mental number line, because improvements transfer to other types of tasks. In interventions aimed at improving children's number line estimates (Opfer and Siegler, 2007; Opfer and Thompson, 2008; Thompson and Opfer, 2008, 2016), children were provided with corrective feedback about the location of the number 150 on a 0–1,000 number line. The feedback alerted children to the fact that their estimates were quite far from the correct location of 150 on the number line. Subsequently, the children scaled their estimates across the entire 0–1,000 numeric range based on their new knowledge of the correct location for 150. To investigate the robustness of this newly adopted linear representation, children were presented with a magnitude categorization transfer task (Opfer and Thompson, 2008). In the magnitude categorization task, five boxes were arranged from left-to-right with a box labeled “really small” for numbers like 0 on the far left and a box labeled “really big” for numbers like 1,000 on the far right. Interestingly, children who made a linear series of number line estimates also made a linear series of category judgments, and this suggests that the linear representation had transferred from one numerical context to another. The left-to-right orientation of the number line and categorization task was aligned with the left-to-right orientation of children's mental number line.

In addition to intervening more directly by providing one-on-one feedback on children's estimates on the number line task, a variety of interventions have aimed to improve children's number sense in more ecologically natural contexts (e.g., board games). There has been increasing interest in improving mathematics performance in early school years by improving children's number sense, through formal and informal instruction (Berkowitz et al., 2015; Ramani and Siegler, 2015; Fazio et al., 2016; Hamdan and Gunderson, 2017). These interventions suggest that learning is improved when the

affordances of materials are aligned with the properties of the mental number line (Siegler, 2016). Next, we review recent research on the use of board games to improve children's mathematics performance.

Interventions using board games have demonstrated learning benefits for children (Ramani and Siegler, 2008; Siegler and Ramani, 2008, 2009; Whyte and Bull, 2008; Ramani et al., 2012; Laski and Siegler, 2014). Board games provide overlapping cues for children to learn the relations between number words and their relative magnitudes (e.g., moving ten spaces from left-to-right takes the child more time to execute and a larger number of moves than moving two spaces from left-to-right). Children who played a board game with ten numbered spaces oriented from left-to-right made larger learning gains than children who played an analogous color board game without consecutively numbered spaces. Specifically, playing the board game for four 15 min sessions, in which smaller numbers were presented in spaces on the left and larger numbers were presented in spaces on the right, improved children's numeral identification, number line estimation, and magnitude comparison performance (Ramani and Siegler, 2008; Siegler and Ramani, 2008). A subsequent experiment investigated the role of linearity in supporting learning by comparing the effects of a linear board game (i.e., spaces numbered 1–10 in a left-to-right orientation) and a circular board game (i.e., spaces numbered 1–10 in a clockwise orientation) (Siegler and Ramani, 2009). The results demonstrated larger learning gains for children who played the linear game because it was hypothesized that the linear board game was better aligned with children's mental number line as compared to the circular board game.

Evidence from these board game experiments suggests that three affordances appear to be most important for materials that support learning number magnitudes: (1) left-to-right orientations, (2) promoting equal spacing between values, and (3) providing an explicit means for mapping numbers to relative magnitudes (Siegler and Booth, 2004; Whyte and Bull, 2008; Laski and Siegler, 2014; Ramani and Siegler, 2015).

Spatial-Numeric Affordances of Books

Reading books to children is an important aspect of promoting children's developing literacy. Sharing reading with young children promotes an understanding of reading conventions (e.g., orthography oriented from left-to-right and top-to-bottom; Whitehurst and Lonigan, 1998) and introduces children to skills related to later reading (e.g., phonemic awareness; Justice et al., 2005). Discussions during shared reading that prompt children to make inferences beyond text improve children's vocabulary and comprehension (Zucker et al., 2013). Books not only support children's developing literacy, but support their developing numeracy. Books can provide support for number and math learning by providing content (e.g., novel words) and opportunities for social interactions (Montag et al., 2015), learning number words (Ward et al., 2017), providing practice for number skills (Skwarchuk et al., 2014), learning relational quantity words like "equal, more, or less" (Hassinger-Das et al., 2015), and improving spatial reasoning (Gunderson et al., 2012). One heretofore unexamined dimension is that the affordances of

the book may provide supports for spatial-numeric learning for relative magnitudes, much like number lines and board games.

The affordances of books may be analogous to number lines and linear board games because they provide overlapping cues for mapping number words to approximate magnitudes (see **Table 1** for comparisons). Recall that the left-to-right orientation of number lines and linear board games was related to greater increases in learning. Books are oriented left-to-right in a similar fashion with smaller page numbers on the left and increasingly larger page numbers on the right. Number lines promote equal spacing between values because the distance between end markers can be evenly divided by equally spaced hatch marks (see Siegler and Opfer, 2003; Schneider et al., 2008; Siegler et al., 2011; and Ashcraft and Moore, 2012, for children's spontaneous segmentation of number lines and Siegler and Thompson, 2014; Peeters et al., 2017a,b for children's use of experimenter-imposed landmarks as they estimated numbers on number lines). Linear board games are structured such that each space represents one value, and moves between spaces are all the same distance. Books have similar affordances in that each page contains two numbers, one on the front and one on the back of each page, and each page turn moves the same distance between the first and last page. Finally, number lines and linear board games provide a means for helping children map numbers to relative magnitudes.

Current Study

In our current study, we created a novel page finder magnitude estimation task in which we asked children to find pages in a 100-page book that did not include printed page numbers. We anticipated that number line estimation performance in the 0–100 range would be related to performance on this page finder task because we oriented children to the book by verbally labeling the first six pages. For this reason, we expected that children might draw comparisons between the 23 cm wide number lines and the 1 cm wide book to decide that the book was simply a smaller, scaled-down version of the number line that did not include printed numeric labels. The classic literature on scale errors suggests that it is not uncommon for preschoolers to attempt to interact with small-scale objects (e.g., tiny replica of a car) in much the same way that they previously interacted with large-scale objects (e.g., large car) (DeLoache et al., 2004). Further, in the domain of mathematics, even infants and young children who do not have formal multiplication and division experience, can perform multiplicative scaling in a non-symbolic context (McCrink and Wynn, 2007; McCrink and Spelke, 2010, 2016). Finally, children transfer their knowledge of linearly arranged numbers to other non-numeric stimuli, such as their estimates of the locations of letters of the alphabet on an ABC line (Hurst et al., 2014).

Our sixty first grade participants completed five tasks in a counterbalanced order: number line estimation (e.g., Where does 25 go on a line with left endpoint labeled 0 and right endpoint labeled 100?), magnitude comparison (e.g., Which is bigger 89 or 54?), numeral identification (e.g., What number is this: 17?), counting on (e.g., Can you count up five more from 84?), and a page finder magnitude estimation task (e.g., Can you find page 33?). We hypothesized that: (1) Magnitude comparison, numeral

identification, and counting on performance would be correlated with number line performance because all of these tasks tap numerical knowledge, and (2) To the extent that the page finder magnitude estimation task also taps magnitude understanding, number line estimation performance will predict page finder performance, even after controlling for age and accuracy on the magnitude comparison, numeral identification, and counting on tasks.

MATERIALS AND METHODS

Participants

Participants were 60 first grade students (M age = 6.68, SD = 0.89) in four classrooms in two public school districts in northeast Ohio. Approximately 39% of children who attended these schools were eligible for free or reduced price lunches. Gender was balanced in the sample: 50% of children were identified as female. Parents provided written informed consent for their children to participate, and children provided verbal assent. Each child received a sticker at the end of the experimental session. The Kent State University IRB approved this study.

Tasks and Procedure

Participants completed five tasks: number line estimation, numeral identification, counting on, magnitude comparison, and page finder magnitude estimation. The number line estimation task is a measure that assesses children's magnitude knowledge; numeral identification is a task that measures children's ability to verbally identify numbers in the 0–100 range that were presented in the other numerical tasks such as number line estimation and the novel page finder task; counting on is a measure that assesses children's numerical knowledge such as the ability to make decade changes as they count; magnitude comparison is a measurement that assesses children's ability to compare numbers in the 0–100 range, and we believed this would be important as children compared the current and previous pages that they found in the page finder task (e.g., "I just found page ____, and now I have to find a bigger page number, page ____.") Task order was counterbalanced, and the problems were presented in a random order within each task. All children were tested individually in a quiet location in their school by a female research assistant.

Number Line Estimation

Children estimated the location of 24 numbers on 23 cm number lines. The lines had a 0 at the left endpoint and a 100 at the right endpoint. One to-be-estimated number appeared at the top left of each page. Children indicated the location of this number by making a vertical hatch mark through the line. When children finished making each estimate, the page was turned over so that they could no longer reference their answer. All children were first asked to point to the location of 0 and 100 and were provided with corrective feedback if they did not point to the correct locations. Then, they estimated the following numbers, that spanned the entire 0–100 range, without feedback from the researcher: 3, 4, 6, 8, 12, 17, 21, 23, 25, 29, 33, 39, 43, 48, 52,

57, 61, 64, 72, 79, 81, 84, 90, and 96. This set of numbers oversamples children's estimates at the low end of the numerical range, consistent with prior research (Opfer et al., 2016). In line with prior research (e.g., Siegler and Booth, 2004; Laski and Siegler, 2007; Booth and Siegler, 2008; Laski and Siegler, 2014), we assessed three aspects of children's estimates: their PAE, the linearity of children's estimates, and the slope of their best fitting linear function. PAE is the absolute difference between the child's estimate and the actual location of the number divided by the scale and expressed as a percentage (i.e., multiplied by 100). Smaller PAE indicates a more accurate series of estimates. Linearity and slope are calculated by regressing each child's set of estimates on the true magnitude of the given numbers. The R^2_{Lin} represents the percent of variance in each child's estimates accounted for by the best fitting linear model for that child. The slopes (b_j) of the best fitting linear model for each child indexes how close that child's estimates are to the ideal slope that relates estimates to the given numbers (1.00). It should be noted that we chose to characterize children's estimates with a linear function to maintain consistency with prior research on informal tasks (i.e., board games, Siegler and Ramani, 2008) associated with children's number line estimates, however, there are other statistical methods for characterizing children's behavior on this task (e.g., Barth and Paladino, 2011; see Opfer et al., 2016 for a discussion).

Numeral Identification

Children named 24 numbers, one at a time, as they were presented on a computer screen. The numbers were the same as those from the number line estimation task. The dependent variable was percentage correct out of 24 trials.

Counting On

This game was adapted from Laski and Siegler (2014) because "counting on" has been established as an important aspect of the typical numerical board game procedure (e.g., when children are on space 5, and they spin a 2, they must say, "6, 7" instead of "1, 2"). Children heard a number (7, 18, 37, and 84), and they were asked to count up by 3, 5, and 8 from each of those starting numbers (e.g., "7, 8, 9, 10, 11, 12"). They were first given the sample problem, "If I say, 'Start counting with one and count up two more numbers,' you would say, '1, 2, 3.'" To ease the working memory burden of the task, children were presented with a linear array of counting chips that corresponded to the number that they had to count up. They were shown the strategy of pointing to each chip as they counted, and they were reminded that they should say the first number and then point to each chip once as they said the next number in the sequence. The dependent variable was percentage correct out of 12 trials. The child could not make any counting errors on a trial for it to be counted as correct.

Magnitude Comparison

Participants were told that they would see two numbers between 0 and 100, and they should compare the numbers to decide which one was bigger. All comparisons contained the number 54, which was chosen because it is close to the midpoint of the

0–100 range (see Siegler et al., 2011 for a similar methodology used in a fraction magnitude comparison task). It was assumed that if children were asked to compare all numbers to 50, this would make the task too easy and would also provide unintended clues about the midpoint of the 0–100 numerical range. Then, children could potentially use these clues as feedback to improve their number line estimation performance (see Opfer et al., 2016). The following numbers were compared with 54: 2, 8, 12, 26, 34, 42, 67, 73, 89, 97. In half of the trials, 54 appeared on the left side of the screen, and in the other half of trials 54 appeared on the right side of the screen. The dependent variable was percentage correct out of 20 trials.

Page Finder Magnitude Estimation

Children were presented with a 100-page book. The book did not include any page numbers. The front and back cover of the book was white, and the book was spiral bound. The children were told that they were going to play a search game. The researcher said a number, and the child was instructed to flip to that page without counting. The researcher said, “Just like one of your books at home, Page 1 is on this side (researcher pointed to page), and Page 2 is on this side (researcher flipped the page and pointed to it). If Page 3 is on this side (researcher pointed to it), which page is on this side (researcher flipped the page and pointed to it)? When children answered correctly, they were told, “Good!” When children answered incorrectly, they were told, “It would be Page 4, right?”

The researcher continued with the instructions, “If this is Page 5 (researcher pointed to page) which page is on this side (researcher flipped the page and pointed to it)?” Again, children were given corrective feedback on this practice trial (i.e., “Good,” or “It would be Page 6, right?”). Then, the child was told, “The book keeps going until we get to page 100 (researcher flipped to page 100).” Then, the child was asked to find page 1 and page 100, and they were given corrective feedback if they did not correctly identify these practice pages.

Children did not receive any corrective feedback on the remaining test trials. They were told, “If I say the number ‘20,’ I want you to quickly flip the pages until you believe that you’ve gotten to page 20. See you can quickly flip through the pages like this.” The researcher demonstrated how to quickly

fan through the pages. Children were reminded how to properly flip through the book if they attempted to count the pages. This most frequently happened when they were asked to find a small number page. It should be noted that some children chose to flip from the back of the book or lift a chunk of pages when the book was closed to get closer to the intended location of a large-numbered page in the book. According to our research assistant, flipping from the back of the book was rare, though admissible in our protocol.¹ After the child flipped to the intended page, he was asked to find a hidden picture on the page. The researcher closed the book before the child searched for the location of the next page. Some children used the strategy of lifting a large chunk of pages to get to the back of the book if asked to find a large number page. We did not systematically code children’s flipping strategies for later analysis.

If children forgot the number of the page that they were looking for, the researcher could verbally remind them by saying, “Where is page N?” Children were asked to search for the following pages: 4, 8, 17, 23, 29, 33, 48, 57, 61, 72, 84, and 90. In line with the number line estimation task, we calculated percent absolute error, PAE = (|page number that the child flipped to – the actual page number| / 100) * 100, linearity (R^2_{Lin}), and slopes of the best fitting linear models.

RESULTS

First, we examined children’s average performance on each task, see **Table 2**. As shown in **Table 2**, children had high accuracy on the numeral identification, magnitude comparison, and counting on tasks, indicating knowledge of numerical symbols. Furthermore, on average, children’s number line estimates were moderately good, with an average PAE of 14%. However, there was substantial variability in the accuracy and precision of children’s number line estimates.

Importantly, children’s performance on most of these tasks was in line with findings from prior research using these tasks with similar age groups. As shown in **Table 2**, children’s

¹We would like to thank a reviewer for suggesting that flipping from the back of the book is parallel to estimating from the right (large) endpoint on the number line estimation task.

TABLE 2 | Average performance in current and prior studies.

Task	Mean: current study	Mean: prior studies	Prior study citation
Number line estimation: PAE	18% (7%)	18% (1st) and 14% (1st); 13% (1st)	Siegler and Booth, 2004; Laski and Siegler, 2007
Number line estimation: R^2_{Lin}	0.52 (0.28)	0.49 (K), 0.90 (1st)	Siegler and Booth, 2004
Number line estimation: Slope	0.51 (0.23)	0.33 (K), 0.58 (1st)	Siegler and Booth, 2004
Numeral identification accuracy	96% (6%)	82% (K)	Laski and Siegler, 2014
Counting on accuracy	79% (27%)	18% (K)	Laski and Siegler, 2014
Magnitude comparison accuracy	94% (10%)	95% (1st)	Laski and Siegler, 2007
Page finder magnitude estimation: PAE	16% (5%)	N/A	N/A
Page finder magnitude estimation: R^2_{Lin}	0.54 (0.26)	N/A	N/A
Page finder magnitude estimation: Slope	0.59 (0.23)	N/A	N/A

Standard deviations reported in parentheses along with corresponding means. Sample characteristics from prior research are reported in parentheses along with corresponding means; K, kindergarten; 1st, first grade.

TABLE 3 | Pairwise correlations between tasks.

Task	Numeral identification	Counting on	Magnitude comparison	Page finder PAE
Number line estimation PAE	−0.49**	−0.38**	−0.57**	0.39**
Numeral identification accuracy		0.30*	0.68**	−0.18
Counting on accuracy			0.58**	−0.07
Magnitude comparison accuracy				−0.27*

Single asterisks denote correlations significant at $p < 0.05$. Double asterisks denote correlations significant at $p < 0.01$. For correlations with Magnitude Comparison, $n = 59$, otherwise, $n = 60$.

accuracy on numeral identification and magnitude comparison was consistent with prior research with first graders and kindergartners (Laski and Siegler, 2007; Laski and Siegler, 2014). In light of the replicability crisis in psychology (Open Science Collaboration, 2015), we wanted to show that our results were consistent with the existing numerical cognition literature. Note that data were collected from our first grade participants in the early part of the academic year (i.e., October and November), and it is for this reason that their performance on some tasks may resemble that of kindergartners from the previous literature. Furthermore, children's average error (PAE), linearity (R^2_{Lin}), and slopes on the number line task were also consistent with prior research (Siegler and Booth, 2004; Laski and Siegler, 2007, 2014). In contrast to prior literature, the children in our sample were more accurate on the counting on task compared to prior research demonstrating poor counting on performance among kindergartners (Laski and Siegler, 2014). Knowledge of the number system develops rapidly across kindergarten and

first grade, and thus this difference in performance may reflect differences in the timing of data collection across the current study and prior research.

Second, we tested for correlations between accuracy on all pairs of tasks. Consistent with prior literature (Laski and Siegler, 2007; Ramani and Siegler, 2008; Siegler and Ramani, 2008, 2009), we expected that children's accuracy on the numeral identification, counting on, and magnitude comparison tasks should be significantly correlated with children's PAE on the number line estimation task. Indeed, this was the case, see **Table 3**. Across all three numeric tasks, lower PAE on the number line estimation task was associated with higher accuracy on the numeric tasks. In other words, as expected, children with more precise representations of whole number magnitude were also more likely to be adept at identifying printed numerals, counting up from a given number, and choosing the larger of two given numbers. Importantly, children's PAE on the number line estimation task was also significantly correlated with PAE on our novel, page finder magnitude estimation task, $r = 0.39$, $p < 0.01$. Children's PAE on the page finder task was also correlated with magnitude comparison, $r = -0.27$, $p = 0.04$, but not significantly correlated with the other numerical tasks that do not measure magnitude knowledge.

Given that the precision of children's magnitude estimates during the number line task was highly correlated with the precision of children's magnitude estimates during the page finder estimation task, we assessed whether the linearity (R^2_{Lin}) of their magnitude estimates and the slope of the best fitting lines were also similar across tasks. Both R^2_{Lin} , $r = 0.46$, $p < 0.001$, and slope, $r = 0.42$, $p < 0.001$, were correlated across tasks (see **Figure 1**). This is further evidence that children who made highly linear estimates on the number line were also likely to make highly linear estimates when seeking page numbers in a book.

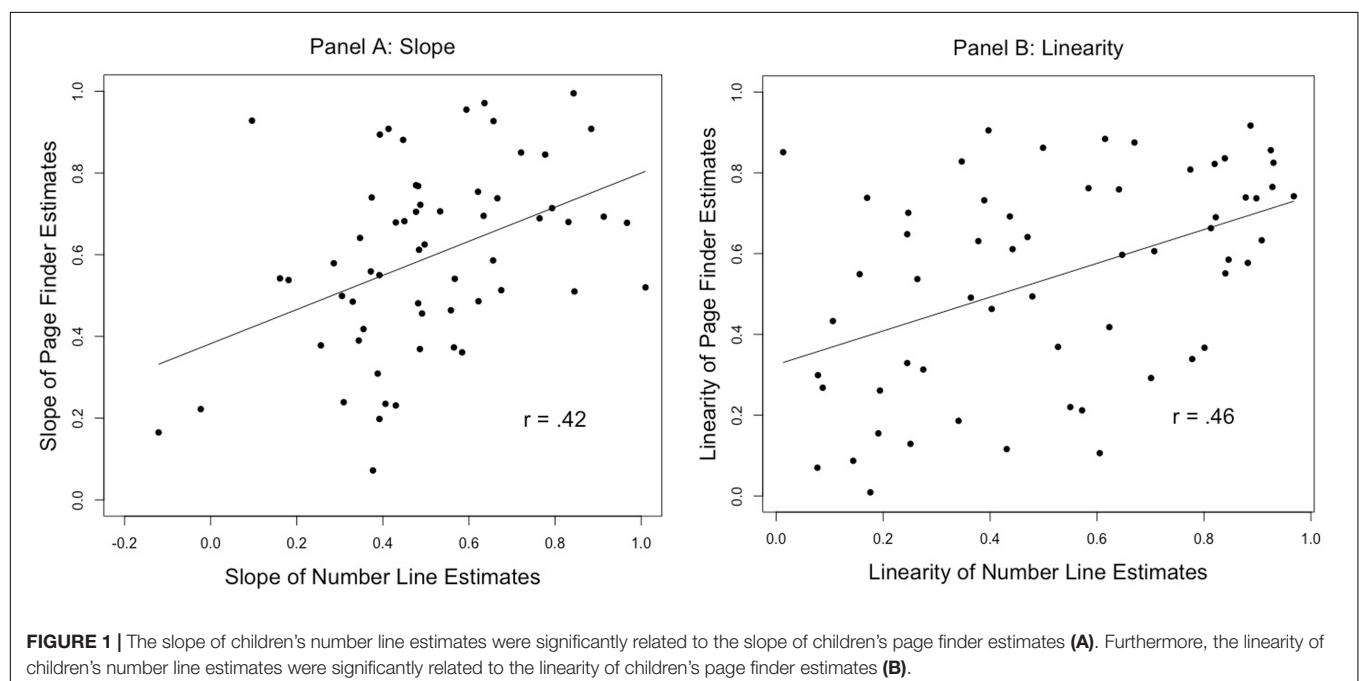


FIGURE 1 | The slope of children's number line estimates were significantly related to the slope of children's page finder estimates (A). Furthermore, the linearity of children's number line estimates were significantly related to the linearity of children's page finder estimates (B).

Finally, we examined whether children's magnitude estimation performance on the page finder task was related to their magnitude estimation performance on the number line estimation task, over and above the other facets of children's numerical knowledge. Although all of the tasks assess children's number knowledge, we hypothesized that the number line estimation task and the page finder magnitude estimation task would both specifically assess magnitude understanding, and therefore would be significantly related even after accounting for other aspects of children's number knowledge. Thus, we regressed children's PAE on the page finder task on children's PAE on the number line task, controlling for accuracy on numeral identification, counting on, and magnitude comparison as well as age. In this model, children's number line estimation PAE did predict their page finder PAE, $\beta = 0.48$, $p < 0.01$, $\eta_p^2 = 0.14$. In contrast, numeral identification, $p = 0.41$, $\eta_p^2 = 0.01$, counting on, $p = 0.20$, $\eta_p^2 = 0.03$, and magnitude comparison, $p = 0.21$, $\eta_p^2 = 0.03$, did not predict children's PAE on the page finder task in this model.

DISCUSSION

Our results provided evidence for a novel measure of spatial-numerical association, the page finder task. We found that for sixty first grade students, their performance on a number line estimation task in the 0–100 range was correlated with their performance on other numerical tasks, such as magnitude comparison, numeral identification, and counting on from a given number. Importantly, all three dependent variables that characterized performance in the number line estimation task (i.e., PAE, R^2_{Lin} , and slope) were related to the same dependent variables in the novel, page finder magnitude estimation task in which children were asked to find the location of a page number in a book. Interestingly, page finder PAE did not correlate with children's accuracy on identifying numerals and counting on from a given number—tasks that seem to rely less on *magnitude* knowledge and more on *symbolic* numerical knowledge—and this may be related to the non-symbolic nature of the page finder book because it contained no printed page numbers. Children's performance on the number line estimation task predicted their page finder PAE, even after controlling for overall age and performance on all other tasks tapping numerical knowledge. Overall, these findings suggest that children may be relying on similar mental representations to guide their estimates on both the highly symbolic number line estimation task and our novel page finder magnitude estimation task that contained no printed numbers.

It was somewhat surprising that children were just as accurate (i.e., similar PAEs and SDs) at finding page numbers in a book without printed page numbers as placing numbers on number lines. The number line estimation task can test for spatial-numeric associations because this task inherently involves spatial (e.g., identifying the physical location of a number on a number line as a distance between the left and right endpoints) as well as numeric components (e.g., end points on the number line, to-be-estimated numbers). Children's accuracy on the page finder

task was all the more impressive because the number line was 23 cm wide, yet the book used in the page finder task was only about 1 cm wide. We interpret children's similar level of accuracy on these tasks as indicating that the number line estimation task and the page finder task tap a common underlying numerical representation. In this way, PAE on each task might indicate the level of precision in the underlying numerical representation: if participants' numerical representations are precise enough to be accurate on one task, they are equally precise and accurate on the other. Was children's performance so accurate on the page finder task because we oriented them to the size of one unit—a procedure similar to that used when children make estimates on “unbounded” number line tasks (Cohen and Sarnecka, 2014)—by orienting them to the first six pages in the book to make sure they understood the task instructions? Similarly, in the zips task (Booth and Siegler, 2006; Thompson and Siegler, 2010), children were shown the length of a 1-unit line and the length of a 1,000-unit line and asked to produce a line of X units. Performance on the zips task correlates with performance on the number line estimation task and a numerosity estimation task in which children fill a jar with a specified number of dots. Performance on these production tasks, such as the page finder, zips, and jar tasks, may all tap children's underlying numerical magnitude representations, much like the number line estimation task.

In our regression analysis, we were able to predict page finder PAE from number line estimation PAE after controlling for age and performance on other numerical tasks, and we take this as evidence in support of the hypothesis that the page finder task and the number line estimation task tap a common underlying numerical representation. It is important to note that we are not able to make any causal claim about the direction of this relationship. In this analysis, we operationalized children's underlying numerical representation by measuring their percent absolute error on the number line estimation task. Thus, we argue that our findings demonstrate that finding a page number in a book taps children's underlying magnitude representation. If this is the case, it may be possible that finding page numbers in books is one way in which parents can help children improve their understanding of relative numerical magnitudes. Parents and teachers already encourage children's literacy development through reading, and reading books is a familiar activity for many children. Our findings suggest that while reading books, caregivers can help children identify page numbers in the books in an effort to promote their understanding of numerical magnitudes. Like board game interventions, books may have the potential to provide an easy and cost-effective means for caregivers to integrate numerical experience into children's everyday lives. In this way, books can promote the development of literacy as well as numeracy skills.

CONCLUSION

Number sense is inherently spatial and numeric (Mix et al., 2016; Leibovich et al., 2017). We investigated whether books share similar spatial-numeric properties of materials, such as

number lines, by using a novel measure, the page finder task. Our results demonstrated strong similarities between page finder estimates and children's number line estimates, which is particularly impressive given that the page finder book was quite small (approximately 1 cm wide) in comparison to the number line. The findings demonstrate the utility of this novel measure and suggest that books share properties with other materials that measure, and potentially improve, children's numerical magnitude knowledge.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the Kent State University IRB with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the Kent State University IRB.

AUTHOR CONTRIBUTIONS

CT and BM conceptualized the study. CT and BM oversaw the data collection with undergraduate research assistants. CT and

PS analyzed the data. CT, BM, and PS wrote and revised the paper.

FUNDING

This work was supported in part by the Institute of Education Sciences, US Department of Education, through Grant R305A160295.

ACKNOWLEDGMENTS

We would like to thank the children and first grade teachers at Longcoy Elementary in the Kent City School District and Wait Primary in the Streetsboro School District in Ohio. We would also like to thank Melissa Bright, Morgan Buerke, Alanna Feltner, Jessica Kukura, Carly Nelson, Rowan Reed, Allison Smolarski, and Jennifer Wagner for their help with stimuli creation, data collection, and data entry. Finally, we would like to acknowledge the feedback we received from Shannon Zentall on an earlier draft of this manuscript.

REFERENCES

- Adachi, I. (2014). Spontaneous spatial mapping of learned sequence in chimpanzees: evidence for a SNARC-like effect. *PLOS ONE* 9:e90373. doi: 10.1371/journal.pone.0090373
- Ashcraft, M. H., and Moore, A. M. (2012). Cognitive processes of numerical estimation in children. *J. Exp. Child Psychol.* 111, 246–267. doi: 10.1016/j.jecp.2011.08.005
- Barth, H. C., and Paladino, A. M. (2011). The development of numerical estimation: evidence against a representational shift. *Dev. Sci.* 14, 125–135. doi: 10.1111/j.1467-7687.2010.00962.x
- Berkowitz, T., Schaeffer, M. W., Maloney, E. A., Peterson, L., Gregor, C., Levine, S. C., et al. (2015). Math at home adds up to achievement in school. *Science* 350, 196–198. doi: 10.1126/science.aac7427
- Booth, J. L., and Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Dev. Psychol.* 42, 189–201. doi: 10.1037/0012-1649.41.6.189
- Booth, J. L., and Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Dev.* 79, 1016–1031. doi: 10.1111/j.1467-8624.2008.01173.x
- Cohen, D. J., and Sarnecka, B. W. (2014). Children's number line estimation shows development of measurement skills not number representations. *Dev. Psychol.* 50, 1640–1652. doi: 10.1037/a0035901
- Dehaene, S. (2011). *The Number Sense*, 2nd Edn. New York, NY: Oxford University Press.
- Dehaene, S., Bossini, S., and Giraux, P. (1993). The mental representation of parity and number magnitude. *J. Exp. Psychol. Gen.* 122, 371–396. doi: 10.1037/0096-3445.122.3.371
- DeLoache, J. S., Uttal, D. H., and Rosengren, K. S. (2004). Scale errors offer evidence for a perception-action dissociation early in life. *Science* 304, 1027–1029. doi: 10.1126/science.1093567
- Fazio, L. K., Bailey, D. H., Thompson, C. A., and Siegler, R. S. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement. *J. Exp. Child Psychol.* 123, 53–72. doi: 10.1016/j.jecp.2014.01.013
- Fazio, L. K., Kennedy, C. A., and Siegler, R. S. (2016). Improving children's knowledge of fraction magnitudes. *PLOS ONE* 11:e0165243. doi: 10.1371/journal.pone.0165243
- Fias, W. (2001). Two routes for the processing of verbal numbers: evidence from the SNARC effect. *Psychol. Res.* 65, 250–259. doi: 10.1007/s004260100065
- Friso-van den Bos, I., Kroesbergen, E., and van Luit, J. E. H. (2014). Number sense in kindergarten children: factor structure and working memory predictors. *Learn. Individ. Dif.* 33, 23–29. doi: 10.1016/j.lindif.2014.05.003
- Gunderson, E. A., Ramirez, G., Beilock, S. L., and Levine, S. C. (2012). The relation between spatial skills and early number knowledge: the role of the linear number line. *Dev. Psychol.* 48, 1229–1241. doi: 10.1037/a0027433
- Hamdan, N., and Gunderson, E. A. (2017). The number line is a critical spatial-numerical representation: evidence from a fraction intervention. *Dev. Psychol.* 53, 587–596. doi: 10.1037/dev0000252
- Hassinger-Das, B., Jordan, N. C., and Dyson, N. (2015). Reading stories to learn math: mathematics vocabulary instruction for children with early numeracy difficulties. *Elem. Sch. J.* 116, 242–264. doi: 10.1086/683986
- Hurst, M., Monahan, K. L., Heller, E., and Cordes, S. (2014). 123s and ABCs: developmental shifts in logarithmic-to-linear responding reflect fluency with sequence values. *Dev. Sci.* 17, 892–904. doi: 10.1111/desc.12165
- Justice, L. M., Skibbe, L., Canning, A., and Lankford, C. (2005). Pre-schoolers, print and storybooks: an observational study using eye movement analysis. *J. Res. Read.* 28, 229–242. doi: 10.1111/j.1467-9817.2005.00267.x
- Landy, D., Silbert, N., and Goldin, A. (2013). Estimating large numbers. *Cogn. Sci.* 37, 775–799. doi: 10.1111/cogs.12028
- Laski, E. V., and Siegler, R. S. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child Dev.* 78, 1723–1743. doi: 10.1111/j.1467-8624.2007.01087.x
- Laski, E. V., and Siegler, R. S. (2014). Learning from number board games: you learn what you encode. *Dev. Psychol.* 50, 853–864. doi: 10.1037/a0034321
- Leibovich, T., Katzin, N., Harel, M., and Henik, A. (2017). From 'sense of number' to 'sense of magnitude' – The role of continuous magnitudes in numerical cognition. *Behav. Brain Sci.* 40:e164. doi: 10.1017/S0140525X16000960
- McCrink, K., and Spelke, E. S. (2010). Core multiplication in childhood. *Cognition* 116, 204–216. doi: 10.1016/j.cognition.2010.05.003
- McCrink, K., and Spelke, E. S. (2016). Non-symbolic division in childhood. *J. Exp. Child Psychol.* 142, 66–82. doi: 10.1016/j.jecp.2015.09.015

- McCrink, K., and Wynn, K. (2007). Ratio abstraction by 6-month-old infants. *Psychol. Sci.* 18, 740–745. doi: 10.1111/j.1467-9280.2007.01969.x
- McCrink, K. M., and Opfer, J. E. (2014). Development of spatial-numerical associations. *Curr. Dir. Psychol. Sci.* 23, 439–445. doi: 10.1177/0963721414549751
- Mix, K. S., Levine, S. C., and Newcombe, N. (2016). “Development of quantitative thinking across correlated dimensions,” in *Continuous Issues in Numerical Cognition: How Many or How Much*, ed. A. Henik (London: Elsevier Inc.), 3–27.
- Montag, J. L., Jones, M. N., and Smith, L. B. (2015). The words children hear picture books and the statistics for language learning. *Psychol. Sci.* 26, 1489–1496. doi: 10.1177/0956797615594361
- Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science* 349, 943–951. doi: 10.1126/science.aac4716
- Opfer, J. E., and Siegler, R. S. (2007). Representational change and children’s numerical estimation. *Cogn. Psychol.* 55, 169–195. doi: 10.1016/j.cogpsych.2006.09.002
- Opfer, J. E., Siegler, R. S., and Young, C. J. (2011). The powers of noise-fitting: reply to Barth and Paladino. *Dev. Sci.* 14, 1194–1204. doi: 10.1111/j.1467-7687.2011.01070.x
- Opfer, J. E., and Thompson, C. A. (2008). The trouble with transfer: insights from microgenetic changes in the representation of numerical magnitude. *Child Dev.* 79, 790–806. doi: 10.1111/j.1467-8624.2008.01158.x
- Opfer, J. E., Thompson, C. A., and Furlong, E. E. (2010). Early development of spatial-numeric associations: evidence from spatial and quantitative performance of preschoolers. *Dev. Sci.* 13, 761–771. doi: 10.1111/j.1467-7687.2009.00934.x
- Opfer, J. E., Thompson, C. A., and Kim, D. (2016). Free versus anchored numerical estimation: a unified approach. *Cognition* 149, 11–17. doi: 10.1016/j.cognition.2015.11.015
- Patro, K., Nuerk, H. C., Cress, U., and Haman, M. (2014). How number-space relationships are assessed before formal schooling: a taxonomy proposal. *Front. Psychol.* 14:419. doi: 10.3389/fpsyg.2014.00419
- Peeters, D., Sekeris, E., Verschaffel, L., and Luwel, K. (2017a). Evaluating the effect of labeled benchmarks on children’s number line estimation performance and strategy use. *Front. Psychol.* 8:1082. doi: 10.3389/fpsyg.2017.01082
- Peeters, D., Verschaffel, L., and Luwel, K. (2017b). Benchmark-based strategies in whole number line estimation. *Br. J. Psychol.* 108, 668–686. doi: 10.1111/bjop.12233
- Ramani, G. B., and Siegler, R. S. (2008). Promoting broad and stable improvements in low-income children’s numerical knowledge through playing number board games. *Child Dev.* 79, 375–394. doi: 10.1111/j.1467-8624.2007.01131.x
- Ramani, G. B., and Siegler, R. S. (2015). “How informal learning activities can promote children’s numerical knowledge,” in *Oxford Handbook of Mathematical Cognition*, eds R. C. Kadosh and A. Dowker (Oxford: Oxford University Press), 1135–1153.
- Ramani, G. B., Siegler, R. S., and Hitti, A. (2012). Taking it to the classroom: number board games as a small group learning activity. *J. Educ. Psychol.* 104, 661–672. doi: 10.1037/a0028995
- Schneider, M., Heine, A., Thaler, V., Torbeyns, J., De Smedt, B., Verschaffel, L., et al. (2008). A validation of eye movements as a measure of elementary school children’s developing number sense. *Cogn. Dev.* 23, 409–422. doi: 10.1016/j.cogdev.2008.07.002
- Sidney, P. G., Thompson, C. A., Matthews, P. G., and Hubbard, E. M. (2017). From continuous magnitudes to symbolic numbers: the centrality of ratio. *Behav. Brain Sci.* 40, e190. doi: 10.1017/S0140525X16002284
- Siegler, R. S. (2016). Magnitude knowledge: the common core of numerical development. *Dev. Sci.* 19, 341–361. doi: 10.1111/desc.12395
- Siegler, R. S., and Booth, J. L. (2004). Development of numerical estimation in young children. *Child Dev.* 75, 428–444. doi: 10.1111/j.1467-8624.2004.00684.x
- Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., et al. (2012). Early predictors of high school mathematics achievement. *Psychol. Sci.* 23, 691–697. doi: 10.1177/0956797612440101
- Siegler, R. S., and Opfer, J. E. (2003). The development of numerical estimation: evidence for multiple representations of numerical quantity. *Psychol. Sci.* 14, 237–243. doi: 10.1111/1467-9280.02438
- Siegler, R. S., and Ramani, G. B. (2008). Promoting broad and stable improvements in low-income children’s numerical knowledge through playing number board games. *Child Dev.* 79, 375–394. doi: 10.1111/j.1467-8624.2007.01131.x
- Siegler, R. S., and Ramani, G. B. (2009). Playing linear number board games – but not circular ones – improves low-income preschoolers’ numerical understanding. *J. Educ. Psychol.* 101, 545–560. doi: 10.1037/a0014239
- Siegler, R. S., and Thompson, C. A. (2014). Numerical landmarks are useful – Except when they’re not. *J. Exp. Child Psychol.* 120, 39–58. doi: 10.1016/j.jecp.2013.11.014
- Siegler, R. S., Thompson, C. A., and Opfer, J. E. (2009). The logarithmic-to-linear shift: one learning sequence, many tasks, many time scales. *Mind Brain Educ.* 3, 143–150. doi: 10.1111/j.1751-228X.2009.01064.x
- Siegler, R. S., Thompson, C. A., and Schneider, M. (2011). An integrated theory of whole number and fractions development. *Cogn. Psychol.* 62, 273–296. doi: 10.1016/j.cogpsych.2011.03.001
- Skwarchuk, S. L., Sowinski, C., and LeFevre, J. A. (2014). Formal and informal home learning activities in relation to children’s early numeracy and literacy skills: the development of a home numeracy model. *J. Exp. Child Psychol.* 121, 63–84. doi: 10.1016/j.jecp.2013.11.006
- Slusser, E. B., Santiago, R. T., and Barth, H. C. (2013). Developmental change in numerical estimation. *J. Exp. Psychol. Gen.* 142, 193–208. doi: 10.1037/a0028560
- Starr, A., Libertus, M. E., and Brannon, E. M. (2013). Number sense in infancy predicts mathematical abilities in childhood. *Proc. Natl. Acad. Sci. U.S.A.* 110, 18116–18120. doi: 10.1073/pnas.1302751110
- Thompson, C. A., and Opfer, J. E. (2008). Costs and benefits of representational change: effect of context on age and sex differences in magnitude estimation. *J. Exp. Child Psychol.* 101, 20–51. doi: 10.1016/j.jecp.2008.02.003
- Thompson, C. A., and Opfer, J. E. (2010). How 15 hundred is like 15 cherries: effect of progressive alignment on representational changes in numerical cognition. *Child Dev.* 81, 1768–1786. doi: 10.1111/j.1467-8624.2010.01509.x
- Thompson, C. A., and Opfer, J. E. (2016). Learning linear spatial-numeric associations improves accuracy of memory for numbers. *Front. Psychol.* 7:24. doi: 10.3389/fpsyg.2016.00024
- Thompson, C. A., and Siegler, R. S. (2010). Linear numerical magnitude representations aid children’s memory for numbers. *Psychol. Sci.* 21, 1274–1281. doi: 10.1177/0956797610378309
- van Galen, M. S., and Reitsma, P. (2008). Developing access to number magnitude: a study of the SNARC effect in 7-to 9-year-olds. *J. Exp. Child Psychol.* 101, 99–113. doi: 10.1016/j.jecp.2008.05.001
- Wall, J. L., Thompson, C. A., Dunlosky, J., and Merriman, W. E. (2016). Children can accurately monitor and control their number-line estimation performance. *Dev. Psychol.* 52, 1493–1502. doi: 10.1037/dev0000180
- Ward, J. M., Mazzocco, M. M., Bock, A. M., and Prokes, N. A. (2017). Are content and structural features of counting books aligned with research on numeracy development? *Early Child. Res. Q.* 39, 47–63. doi: 10.1016/j.ecresq.2016.10.002
- Whitehurst, G. J., and Lonigan, C. J. (1998). Child development and emerging literacy. *Child Dev.* 69, 848–872. doi: 10.1111/j.1467-8624.1998.tb06247.x
- Whyte, J. C., and Bull, R. (2008). Number games, magnitude representation, and basic number skills in preschoolers. *Dev. Psychol.* 44:588. doi: 10.1037/0012-1649.44.2.588
- Wood, G., Willmes, K., Nuerk, H.-C., and Fischer, M. H. (2008). On the cognitive link between space and number: a meta-analysis of the SNARC effect. *Psychol. Sci. Q.* 50, 489–525.
- Zucker, T. A., Cabell, S. Q., Justice, L. M., Pentimonti, J. M., and Kaderavek, J. N. (2013). The role of frequent, interactive prekindergarten shared reading in the longitudinal development of language and literacy skills. *Dev. Psychol.* 49, 1425–1439. doi: 10.1037/a0030347

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Thompson, Morris and Sidney. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Why Numbers Are Embodied Concepts

Martin H. Fischer*

Cognitive Science, University of Potsdam, Potsdam, Germany

Keywords: arithmetic, numerical cognition, number concepts, embodied cognition, philosophy of science

Number concepts are often thought to be abstractions, for example because the numerosity of sets (e.g., their “three-ness”) is a feature apparently dissociated from the sensory experiences with specific set members, such as their size, shape, or color. In other words, quantity-specific experiences seem to vary arbitrarily when we enumerate three apples, three cars, three people, or three fingers. Hence, Frege (1884) and other logically-minded philosophers considered positive integers as ideal cognitive constructions for enumerative mental operations, removed from contextual constraints, yet preserving precision and generalizing across situations (e.g., arithmetic operations).

Yet, upon closer consideration several sensory and also motor features systematically co-occur with each enumeration we perform; this co-occurrence establishes experiential patterns through which number concepts become embodied as part of their acquisition history (cf. Fischer and Brugger, 2011). I describe here several such systematic co-occurrences and cite supporting evidence. This psycho-logical view of number is not in conflict with but extends purely logical considerations of number concepts as foundations of formal arithmetic, as proposed by Frege (1884).

To-be enumerated objects are usually all simultaneously available to us and thus, by physical necessity, distributed across space because two objects cannot occupy the same place at the same time. Therefore, more objects take up more space and enumerating them invites spatially distributed and temporally extended behaviors; these are sensory cues to number. The systematic directionality of counting behaviors furthermore establishes spatial-numerical associations which, in turn, can be detected with chronometric methods and through behavioral biases (see Fischer and Shaki, 2014; Winter et al., 2015, for reviews).

We know that set members should be aligned or grouped in space to reduce spatial memory load when counting them. We apply verbal sensory-motor routines to establish one-to-one correspondences between objects and number names until each object (or group) was referenced once and the last number name establishes set size or cardinality (e.g., Gelman and Gallistel, 1978). Without such direct referencing of objects through pointing, our eyes and fingers are the universal means of associating body postures (i.e., spatial, visual, kinesthetic, and proprioceptive signals) with number names (Fischer, 2003a; Di Luca and Pesenti, 2011). As a consequence, eye position predicts numerical thoughts (Loetscher et al., 2010), tactile finger stimulation primes number processing and perceiving numbers in turn modulates visual-spatial (Fischer et al., 2003) as well as tactile sensitivity (Tschemtscher et al., 2012; Sixtus et al., 2017; Sixtus et al., in revision). Even when overt finger movements are avoided, we spontaneously generate repetitive upper-body movements to enrich our counting with sensory-motor feedback (Carlson et al., 2007).

Habitually, people raised in Western cultures point at horizontally distributed objects left-to-right and thereby associate increasingly larger number names with increasingly more right-sided actions (Opfer and Furlong, 2011; Shaki et al., 2012). The origin of this cultural bias can be traced to observational learning at pre-school age (Göbel et al., 2018) but might have evolutionary origins (Rugani et al., 2015). In other words, the ubiquitous spatial-numerical association of response codes (SNARC) effect results from preferred sensory-motor habits.

OPEN ACCESS

Edited by:

Krzysztof Cipora,
Universität Tübingen, Germany

Reviewed by:

Mario Bonato,
Università degli Studi di Padova, Italy

*Correspondence:

Martin H. Fischer
martinf@uni-potsdam.de

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 19 October 2017

Accepted: 22 December 2017

Published: 15 January 2018

Citation:

Fischer MH (2018) Why Numbers Are
Embodied Concepts.
Front. Psychol. 8:2347.
doi: 10.3389/fpsyg.2017.02347

Two other signature effects of numerical cognition may also be embodied in origin and not only epiphenomenally so. First, when deciding which of two sets is physically larger performance is governed by Weber's law where the just-noticeable difference between them increases with set size. Moreover, we heuristically expect the larger set to also be the more numerous. This pattern is preserved when we distinguish symbolic quantities (i.e., discriminating two digits' meanings). This so-called *numerical distance effect* (Moyer and Landauer, 1967) suggests that we obligatorily recur to the concrete sensory and motor experiences present when these concepts were acquired.

And secondly, when gathering objects, we extend our sensory-motor experiences from the audio-visual and motor to the haptic domain. As a result, wider grasp apertures prime larger numbers (Andres et al., 2004) and number magnitude in turn biases ongoing visuo-motor control (Fischer, 2003b). Holding object sets also lets us experience positive correlations of numerosity and weight. Thus, systematic multi-modal sensory-motor experiences accompany the use of natural number concepts and pose scaled processing challenges for the cognitive system. This is the embodied foundation of the *numerical size effect*, i.e., the systematic increase in processing costs associated with larger numerosities, capturing most everyday experiences, such as managing to juggle 3 but not 4 balls (Fischer, 2017).

To add things up, three cardinal signatures of numerical cognition, the SNARC effect, the distance effect, and the size effect, might be grounded in sensory-motor experiences and in this sense embodied (for a terminological distinction between "grounded" and "embodied" numerical processing, see Fischer, 2012). It is therefore not surprising that we find cross-domain priming in a wide range of tasks whenever people think quantitatively, be they temporal, spatial, or conceptual (Casasanto and Boroditsky, 2008; Scheepers and Sturt, 2014; Walsh, 2015). These associations extend beyond the positive

integers or their manipulation in mental arithmetic (e.g., Werner and Raab, 2014) and even shape how we think about negative numbers that cannot be experienced as sensory quantities. An initial report (Fischer, 2003b) associated negative numbers with left-sided space and also showed a size effect (while controlling the distance effect). The finding generated some controversy (reviewed in Mende et al., 2017) but was confirmed when the assessment removed potential biases from spatially distributed stimulus presentation or response recording (Fischer and Shaki, 2017). Our habitual experience with spatially organized magnitudes thus replaces the lack of sensory experience with negative numbers to generate predictable sensory-motor associations.

In conclusion, number concepts, although often used in a context-free and seemingly abstract manner (Frege, 1884), always carry sensory-motor connotations. This correlative experience is used for prediction not abstraction—in other words, we apply concrete experiences gathered within a knowledge domain (the source) to generate predictions that enrich seemingly "abstract" conceptual knowledge (the target domain). Thus, it is only through the embodied lens that we can appreciate the full nature of number knowledge and devise appropriate methods for effective training and rehabilitation of numerical cognition.

AUTHOR CONTRIBUTIONS

MF: conceived of and wrote this comment and is solely responsible for the content.

FUNDING

This research was funded by the German research foundation (DFG) FI 1915/5-1 (A motor priming approach to problem solving).

REFERENCES

- Andres, M., Davare, M., Pesenti, M., Olivier, E., and Seron, X. (2004). Number magnitude and grip aperture interaction. *Neuroreport* 15, 2773–2777.
- Carlson, R. A., Avraamides, M. N., Cary, M., and Strasberg, S. (2007). What do the hands externalize in simple arithmetic. *J. Exp. Psychol.* 33, 747–756. doi: 10.1037/0278-7393.33.4.747
- Casasanto, D., and Boroditsky, L. (2008). Time in the mind: using space to think about time. *Cognition* 106, 579–593. doi: 10.1016/j.cognition.2007.03.004
- Di Luca, S., and Pesenti, M. (2011). Finger numeral representations: more than just another symbolic code. *Front. Psychol.* 2:272. doi: 10.3389/fpsyg.2011.00272
- Fischer, M. H. (2003a). Spatial representations in number processing – evidence from a pointing task. *Vis. Cogn.* 10, 493–508. doi: 10.1080/13506280244000186
- Fischer, M. H. (2003b). Cognitive representation of negative numbers. *Psychol. Sci.* 14, 278–282. doi: 10.1111/1467-9280.03435
- Fischer, M. H. (2012). A hierarchical view of grounded, embodied and situated numerical cognition. *Cogn. Process.* 13, S161–S164. doi: 10.1007/s10339-012-0477-5
- Fischer, M. H. (2017). "Why numbers are(n't) abstract concepts," in *Presentation at the 20th Meeting of the European Society for Cognitive Psychology* (Potsdam).
- Fischer, M. H., and Brugger, P. (2011). When digits help digits: spatial-numerical associations point to finger counting as prime example of embodied cognition. *Front. Psychol.* 2:260. doi: 10.3389/fpsyg.2011.00260
- Fischer, M. H., Castel, A. D., Dodd, M. D., and Pratt, J. (2003). Perceiving numbers causes spatial shifts of attention. *Nat. Neurosci.* 6, 555–556. doi: 10.1038/nn1066
- Fischer, M. H., and Shaki, S. (2014). Spatial associations in numerical cognition: from single digits to arithmetic. *Q. J. Exp. Psychol.* 67, 1461–1483. doi: 10.1080/17470218.2014.927515
- Fischer, M. H., and Shaki, S. (2017). Implicit spatial-numerical associations: negative numbers and the role of counting direction. *J. Exp. Psychol.* 43, 639–643. doi: 10.1037/xhp0000369
- Frege, G. (1884). *Die Grundlagen der Arithmetik (Foundations of Arithmetic)*. Breslau: Wilhelm Koebner.
- Gelman, R., and Gallistel, C. R. (1978). *The Child's Understanding of Number*. Cambridge, MA: Harvard University Press.
- Göbel, S. M., McCrink, K., Fischer, M. H., and Shaki, S. (2018). Observation of directional storybook reading influences young children's counting direction. *J. Exp. Child Psychol.* 166, 49–66. doi: 10.1016/j.jecp.2017.08.001
- Loetscher, T., Bockisch, C. J., Nicholls, M. E. R., and Brugger, P. (2010). Eye position predicts what number you have in mind. *Curr. Biol.* 20, R264–R265. doi: 10.1016/j.cub.2010.01.015
- Mende, M. A., Shaki, S., and Fischer, M. H. (2017). On the development and representation of negative number concepts (this special research topic).
- Moyer, R. S., and Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature* 215, 1519–1520. doi: 10.1038/2151519a0
- Opfer, J. E., and Furlong, E. E. (2011). How numbers bias preschoolers' spatial search. *J. Cross Cul. Psychol.* 42, 682–695. doi: 10.1177/0022022111406098

- Rugani, R., Vallortigara, G., Priftis, K., and Regolin, L. (2015). Number-space mapping in the newborn chick resembles humans' number line. *Science* 347, 534–536. doi: 10.1126/science.aaa1379
- Scheepers, C., and Sturt, P. (2014). Bi-directional syntactic priming across cognitive domains: from arithmetic to language and back. *Q. J. Exp. Psychol.* 67, 1643–1653. doi: 10.1080/17470218.2013.873815
- Shaki, S., Fischer, M. H., and Göbel, S. M. (2012). Direction counts: a comparative study of spatially directional counting biases in cultures with different reading directions. *J. Exp. Child Psychol.* 112, 275–281. doi: 10.1016/j.jecp.2011.12.005
- Sixtus, E., Fischer, M. H., and Lindemann, O. (2017). Finger posing primes number comprehension. *Cogn. Process.* 18, 237–248. doi: 10.1007/s10339-017-0804-y
- Tschentscher, N., Hauk, O., Fischer, M. H., and Pulvermüller, F. (2012). You can count on the motor cortex: fMRI reveals embodied number processing. *NeuroImage* 59, 3139–3148. doi: 10.1016/j.neuroimage.2011.11.037
- Walsh, V. (2015). "A theory of magnitude: the parts that sum to number," in *The Oxford Handbook of Numerical Cognition*, eds R. Cohen Kadosh and A. Dowker (Oxford: University Press), 552–565.
- Werner, K., and Raab, M. (2014). Moving your eyes to solution: effects of movements on the perception of a problem-solving task. *Q. J. Exp. Psychol.* 67, 1571–1578. doi: 10.1080/17470218.2014.889723
- Winter, B., Matlock, T., Shaki, S., and Fischer, M. H. (2015). Mental number space in three dimensions. *Neurosci. Biobehav. Rev.* 57, 209–219. doi: 10.1016/j.neubiorev.2015.09.005
- Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Fischer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Commentary: The mental representation of integers: An abstract-to-concrete shift in the understanding of mathematical concepts

Melinda A. Mende¹, Samuel Shaki² and Martin H. Fischer^{1*}

¹ Division of Cognitive Sciences, Department of Psychology, University of Potsdam, Potsdam, Germany, ² Department of Behavioral Sciences, Ariel University, Ariel, Israel

Keywords: cognitive development, mental number line, negative numbers, embodied cognition, abstract concepts

A Commentary on

The mental representation of integers: An abstract-to-concrete shift in the understanding of mathematical concepts

by Varma, S., and Schwartz, D. L. (2011). *Cognition* 121, 363–385. doi: 10.1016/j.cognition.2011.08.005

OPEN ACCESS

Edited by:

Frank Domahs,
Philipps University of Marburg,
Germany

Reviewed by:

Ken Hiraiwa,
Meiji Gakuin University, Japan
Sashank Varma,
University of Minnesota, United States
Daniel L. Schwartz,
Stanford University, United States

*Correspondence:

Martin H. Fischer
martinf@uni-potsdam.de

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 17 November 2017

Accepted: 07 February 2018

Published: 27 February 2018

Citation:

Mende MA, Shaki S and Fischer MH
(2018) Commentary: The mental
representation of integers: An
abstract-to-concrete shift in the
understanding of mathematical
concepts. *Front. Psychol.* 9:209.
doi: 10.3389/fpsyg.2018.00209

Decision times during processing of positive number symbols (1, 2, 3 etc.) inform our understanding of mental representations of integers (Holyoak, 1978; Dehaene et al., 1993; Fischer and Shaki, 2014). Effects of number magnitude on cognition include *distance effects* (faster discrimination for larger numerical differences in a number pair), *size effects* (faster processing of smaller numbers), *Spatial-Numerical Association of Response Codes* (SNARC; faster left/right responses to small/large numbers), linguistic markedness (MARC; faster left/right responses to odd/even numbers) and *semantic congruity effects* (faster smaller/larger decisions over smaller/larger number pairs). Results converge on the notion of a spatially oriented mental number line (MNL) where numerically smaller number concepts exist to the left of larger number concepts. How do these performance signatures help us to understand the cognitive representation of negative number symbols (−1, −2, −3 etc.)? Unlike natural number symbols, negative number symbols lack corresponding real entities that support sensory-motor learning. We discuss a recent proposal by Varma and Schwartz (2011) with implications for developmental research.

TERMINOLOGICAL CLARIFICATION

Different terms distinguish two fundamentally different views regarding the cognitive representation of negative numbers: The first view states that negative numbers are cognitively represented to the left of positive numbers, thereby extending the MNL infinitely leftward (henceforth called “extended MNL account”). The second view states that negative numbers have no cognitive representations but are understood through augmenting positive entries of the MNL (henceforth called “rule-based MNL account”). This dichotomy reflects identical distinctions made by Fischer (2003: ontogenetic vs. phylogenetic), Shaki and Petrusic (2005: extended number line vs. magnitude polarity), Ganor-Stern and Tzelgov (2008: holistic vs. components) and Varma and Schwartz (2011: analog+ vs. symbol+). Evidence from magnitude comparisons was used to support either account (see **Table 1** for more studies) so we review it before recommending methodological improvements.

EVIDENCE FROM MAGNITUDE COMPARISON

Magnitude comparison was first used by Fischer (2003) to report a cognitive processing signature for negative numbers: Adults identified the larger of two digits ranging from -9 to $+9$ and

shown in pairs with constant numerical distance (to control both distance and MARC effects). Faster decisions obtained when the spatial arrangement of digits on screen matched a leftward-extended mental number line, thus supporting the extended MNL account. However, Shaki and Petrusic (2005) identified a confound with semantic congruity and showed that results

TABLE 1 | Summary of previous empirical work on the cognitive representation of negative numbers.

Task	Stimuli	Responses	Measures	References
Magnitude comparison	Spatial	Spatial (discrete)	SNARC effect	Fischer, 2003
Magnitude comparison	Spatial	Spatial (discrete)	SNARC + Semantic congruity effect	Shaki and Petrusic, 2005
Magnitude comparison	Spatial	Spatial (discrete)	Distance effect + Semantic congruity effect	Ganor-Stern, 2012
Magnitude comparison	Spatial	Spatial (discrete)	Distance effect + Semantic congruity effect	Ganor-Stern et al., 2010
Magnitude comparison	Spatial	Spatial (discrete)	Semantic congruity effect + Size effect	Ganor-Stern and Tzelgov, 2008
Magnitude comparison	Spatial	Spatial (discrete)	Distance effect + Semantic congruity effect	Tzelgov et al., 2009
Magnitude comparison	Spatial	Spatial (discrete)	Sign-decade compatibility effect	Huber et al., 2015
Magnitude comparison	Spatial	Spatial (discrete)	Distance effect + Semantic congruity effect + Size effect	Varma and Schwartz, 2011
Magnitude comparison	Spatial	Spatial (discrete)	(Neural) Distance effect (fMRI)	Blair et al., 2012
Magnitude comparison	Spatial	Spatial (discrete)	(Neural) Distance effect (fMRI)	Gullick et al., 2012; Gullick and Wolford, 2013
Physical comparison	Spatial	Spatial (discrete)	Distance effect + Size Congruity Effect	Tzelgov et al., 2009
Physical comparison	Spatial	Spatial (discrete)	Congruity Effect (ERP)	Parnes et al., 2012
Magnitude comparison	Spatial	Spatial (discrete)	Number mining (fMRI)	Chassy and Grodd, 2012
Magnitude classification	Centered	Spatial (discrete)	Distance effect + SNARC	Krajcsi and Igács, 2010
Magnitude classification	Centered	Spatial (discrete)	SNARC effect	Fischer and Rottmann, (2005, Experiment 2)
Parity classification + Priming	Centered	Spatial (discrete)	SNARC effect	Tse and Altarriba, 2010
Parity classification	Centered	Spatial (discrete)	SNARC effect	Fischer and Rottmann (2005, Experiment 1)
Parity classification	Centered	Spatial (discrete)	SNARC effect	Nuerk et al., 2004
Parity classification	Centered	Spatial (discrete)	SNARC effect	Prather and Boroditsky, 2003
Pointing (Number line)	Centered	Spatial (continuous)	Scalar variability model	Ganor-Stern and Tzelgov, 2008
Pointing (Number line)	Centered	Spatial (continuous)	Linear or logarithmic	Young and Booth, 2015
Center Classification (Number line)	Spatial	Verbal	Leftward bias, SNARC effect	Loftus et al. (2009, Experiment 2)
Detection (visual)	Centered digit + Spatial target	Centered	Spatial shift of attention	Dodd, 2011
Detection (visual)	Centered digit + Spatial target	Centered	Spatial shift of attention	Zhang and You, 2012
Detection (auditory)	Centered digit + Spatial target	Centered	Spatial shift of attention	Kong et al., 2012

depend on whether positive and negative numbers are blocked or mixed.

Ganor-Stern and Tzelgov (2008) found similar size effects for positive and negative numbers in the comparison task and a systematic decrease of localization variability with increasing number magnitude in a number-to-position task (where adults localized the position of numbers with a mouse cursor on a horizontal line). They inferred a rule-based MNL account.

Varma and Schwartz (2011) found an inverse distance effect in magnitude comparison with adults, inconsistent with a rule-based MNL which predicted no distance effect at all in mixed comparisons (with one positive and one negative integer), due to superficial sign comparisons. The authors augmented the extended MNL account by postulating additional knowledge about the relationship between positive and negative number concepts which is not available yet to 6th graders because they showed no inverse distance effect and thus used a rule-based MNL.

EVIDENCE FROM OTHER METHODS

This conclusion is surprising, given the wide consensus for a concrete-to-abstract shift in knowledge development. Why are conclusions so heterogeneous, even when using a single task? Other methods assessed negative number representation, including pointing, parity judgments, brain activation, eye movement recording and computer simulation (see **Table 1** for details). For example, Gullick and Wolford (2013) investigated neural distance effects in children. They found that IPS activity increased with age while parietal, frontal and precentral activity decreased, consistent with an anterior-posterior shift during maturation (Rivera et al., 2005). They concluded that practice and experience help to integrate negative numbers into an extended mental number line. In addition, Young and Booth (2015) found results both in line with an extended MNL and in line with a rule-based MNL account in two pointing experiments with middle school students. The authors concluded that this conflicting pattern could reflect under-developed number knowledge and differences in previous number exposure. In summary, previous findings in adult and children studies are highly controversial.

REFERENCES

- Blair, K. P., Rosenberg-Lee, M., Tsang, J. M., Schwartz, D. L., and Menon, V. (2012). Beyond natural numbers: negative number representation in parietal cortex. *Hum. Neurosci.* 6, 1–17. doi: 10.3389/fnhum.2012.00007
- Chassy, P., and Grodd, W. (2012). Comparison of quantities: core and format-dependent regions as revealed by fMRI. *Cereb. Cortex* 22, 1420–1430. doi: 10.1093/cercor/bhr219
- Dehaene, S., Bossini, S., and Giraux, P. (1993). The mental representation of parity and number magnitude. *J. Exper. Psychol.* 122, 371–396. doi: 10.1037/0096-3445.122.3.371
- Dodd, M. D. (2011). Negative numbers eliminate, but do not reverse, the attentional SNARC effect. *Psychol. Res.* 75, 2–9. doi: 10.1007/s00426-010-0283-6
- Fischer, M. H. (2003). Cognitive representation of negative numbers. *Psychol. Sci.* 14, 278–282. doi: 10.1111/1467-9280.03435
- Fischer, M. H., and Rottmann, J. (2005). Do negative numbers have a place on the mental number line? *Psychology Science* 47, 22–32.

The lack of consistent effects in adults does not provide a sufficient basis for firm developmental interpretations, thus distorting current conclusions about the development of negative number processing.

METHODOLOGICAL COMMENT

We believe that this ongoing debate benefits from a methodological comment. Specifically, we note that all published studies on negative number processing either presented spatially distributed stimuli or recorded response speed with lateralized keys (see **Table 1**). This use of spatially distributed stimuli or responses permits participants different strategies (e.g., selective attending to the sign or “mirroring” cf. Varma and Schwartz, 2011) and induces extraneous biases (e.g., the semantic congruity effect), all of which contaminates number processing (Fischer and Rottmann, 2005; Shaki and Petrusic, 2005; Gevers et al., 2010; Fischer and Shaki, 2016).

To address this concern, we recently developed a method where positive and negative numbers are interleaved with spatially oriented objects. Participants only ever see a single stimulus (number or object) and respond with a single button only if the relevant part of a conjunction rule was fulfilled (Fischer and Shaki, 2017). Examples are “respond only if the number is larger than -5 or the car is facing left” (incongruent rule) or “respond only if the number is smaller than -5 or the car is facing left” (congruent rule). We found that negative numbers are associated with space according to their signed magnitude, thus resolving the long-standing debate about the cognitive representation of negative numbers (Fischer, 2003; Shaki and Petrusic, 2005): Once the task prevents strategies, an extended mental number line prevails. This conclusion is based on results from a paradigm free of spatial or reporting biases. It can, in turn, inform our studies of the development of negative number concepts (Shaki and Fischer, 2018).

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

- Fischer, M. H., and Shaki, S. (2014). Spatial associations in numerical cognition—From single digits to arithmetic. *Q. J. Exper. Psychol.* 67, 1461–1483. doi: 10.1080/17470218.2014.927515
- Fischer, M. H., and Shaki, S. (2016). Measuring spatial-numerical associations: evidence for a purely conceptual link. *Psychol. Res.* 80, 109–112. doi: 10.1007/s00426-015-0646-0
- Fischer, M. H., and Shaki, S. (2017). Implicit spatial-numerical associations: negative numbers and the role of counting direction. *J. Exp. Psychol.* 43, 639–643. doi: 10.1037/xhp0000369
- Ganor-Stern, D. (2012). Fractions but not negative numbers are represented on the mental number line. *Acta Psychol.* 139, 350–357. doi: 10.1016/j.actpsy.2011.11.008
- Ganor-Stern, D., Pinhas, M., Kallai, A., and Tzelgov, J. (2010). Holistic representation of negative numbers is formed when needed for the task. *Q. J. Exp. Psychol.* 63, 1969–1981. doi: 10.1080/17470211003721667
- Ganor-Stern, D., and Tzelgov, J. (2008). Negative numbers are generated in the mind. *Exp. Psychol.* 55, 157–163. doi: 10.1027/1618-3169.55.3.157

- Gevers, W., Santens, S., Dhooze, E., Chen, Q., Van den Bossche, L., Fias, W., et al. (2010). Verbal-spatial and visuospatial coding of number-space interactions. *J. Exper. Psychol.* 139, 180. doi: 10.1037/a0017688
- Gullick, M. M., and Wolford, G. (2013). Understanding less than nothing: children's neural response to negative numbers shifts across age and accuracy. *Front. Psychol.* 4, 1–17. doi: 10.3389/fpsyg.2013.00584
- Gullick, M. M., Wolford, G., and Temple, E. (2012). Understanding less than nothing: neural distance effects for negative numbers. *Neuroimage* 62, 542–554. doi: 10.1016/j.neuroimage.2012.04.058
- Holyoak, K. J. (1978). Comparative judgments with numerical reference points. *Cogn. Psychol.* 10, 203–243. doi: 10.1016/0010-0285(78)90014-2
- Huber, S., Cornelsen, S., Moeller, K., and Nuerk, H. C. (2015). Toward a model framework of generalized parallel componential processing of multi-symbol numbers. *J. Exp. Psychol. Learn. Mem. Cogn.* 41, 732–745. doi: 10.1037/xlm0000043
- Kong, F., Zhao, J., and You, X. (2012). Components representation of negative numbers: evidence from auditory stimuli detection and number classification tasks. *Q. J. Exper. Psychol.* 65, 691–701. doi: 10.1080/17470218.2011.622048
- Krajcsi, A., and Igács, J. (2010). Processing negative numbers by transforming negatives to positive range and by sign shortcut. *Eur. J. Cog. Psychol.* 22, 1021–1038. doi: 10.1080/09541440903211113
- Loftus, A. M., Nicholls, M. E., Mattingley, J. B., Chapman, H. L., and Bradshaw, J. L. (2009). Pseudoneglect for the bisection of mental number lines. *Q. J. Exp. Psychol.* 62, 925–945. doi: 10.1080/17470210802305318
- Nuerk, H. C., Iversen, W., and Willmes, K. (2004). Notational modulation of the SNARC and the MARC (linguistic markedness of response codes) effect. *Q. J. Exp. Psychol. Sect. A* 57, 835–863. doi: 10.1080/02724980343000512
- Parnes, M., Berger, A., and Tzelgov, J. (2012). Brain representations of negative numbers. *Can. J. Exper. Psychol.* 66:251. doi: 10.1037/a0028989
- Prather, R. W., and Boroditsky, L. (2003, January). “Left of zero: Representing negative numbers on the mental number line,” in *25th Annual Meeting of the Cognitive Science Society* (Boston, MA).
- Rivera, S. M., Reiss, A. L., Eckert, M. A., and Menon, V. (2005). Developmental changes in mental arithmetic: evidence for increased functional specialization in the left inferior parietal cortex. *Cereb. Cortex* 15, 1779–1790. doi: 10.1093/cercor/bhi055
- Shaki, S., and Fischer, M. H. (2018). Deconstructing spatial-numerical associations. *Cognition*.
- Shaki, S., and Petrusic, W. M. (2005). On the mental representation of negative numbers: context-dependent SNARC effects with comparative judgments. *Psychonomic Bull. Rev.* 12, 931–937. doi: 10.3758/BF03196788
- Tse, C. S., and Altarriba, J. (2010). The revelation of the negative side of a mental number line depends on list context: evidence from a sign priming paradigm. *Eur. J. Cogn. Psychol.* 22, 1248–1260. doi: 10.1080/09541440903427495
- Tzelgov, J., Ganor-Stern, D., and Maymon-Schreiber, K. (2009). The representation of negative numbers: exploring the effects of mode of processing and notation. *Q. J. Exp. Psychol.* 62, 605–624. doi: 10.1080/17470210802034751
- Varma, S., and Schwartz, D. L. (2011). The mental representation of integers: an abstract-to-concrete shift in the understanding of mathematical concepts. *Cognition* 121, 363–385. doi: 10.1016/j.cognition.2011.08.005
- Young, L. K., and Booth, J. L. (2015). Student magnitude knowledge of negative numbers. *J. Num. Cogn.* 1, 38–55. doi: 10.5964/jnc.v1i1.7
- Zhang, Y. and You, X. (2012). Extending the mental number line—How do negative numbers contribute?. *Perception* 41, 1323–1335. doi: 10.1068/p7081

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Mende, Shaki and Fischer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Developmental Changes in the Effect of Active Left and Right Head Rotation on Random Number Generation

Charlotte Sosson[†], Carrie Georges^{*†}, Mathieu Guillaume, Anne-Marie Schuller and Christine Schiltz

Institute of Cognitive Science and Assessment, Research Unit Education, Culture, Cognition and Society, Faculty of Language and Literature, Humanities, Arts and Education, University of Luxembourg, Esch-sur-Alzette, Luxembourg

OPEN ACCESS

Edited by:

Frank Domahs,
Philipps University of Marburg,
Germany

Reviewed by:

Tobias Loetscher,
University of South Australia, Australia
Elena Sixtus,
University of Potsdam, Germany

*Correspondence:

Carrie Georges
carrie.georges@uni.lu

[†]These authors have contributed
equally to this work.

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 17 November 2017

Accepted: 12 February 2018

Published: 28 February 2018

Citation:

Sosson C, Georges C, Guillaume M,
Schuller A-M and Schiltz C (2018)
Developmental Changes in the Effect
of Active Left and Right Head
Rotation on Random Number
Generation. *Front. Psychol.* 9:236.
doi: 10.3389/fpsyg.2018.00236

Numbers are thought to be spatially organized along a left-to-right horizontal axis with small/large numbers on its left/right respectively. Behavioral evidence for this mental number line (MNL) comes from studies showing that the reallocation of spatial attention by active left/right head rotation facilitated the generation of small/large numbers respectively. While spatial biases in random number generation (RNG) during active movement are well established in adults, comparable evidence in children is lacking and it remains unclear whether and how children's access to the MNL is affected by active head rotation. To get a better understanding of the development of embodied number processing, we investigated the effect of active head rotation on the mean of generated numbers as well as the mean difference between each number and its immediately preceding response (the first order difference; FOD) not only in adults ($n = 24$), but also in 7- to 11-year-old elementary school children ($n = 70$). Since the sign and absolute value of FODs carry distinct information regarding spatial attention shifts along the MNL, namely their direction (left/right) and size (narrow/wide) respectively, we additionally assessed the influence of rotation on the total of negative and positive FODs regardless of their numerical values as well as on their absolute values. In line with previous studies, adults produced on average smaller numbers and generated smaller mean FODs during left than right rotation. More concretely, they produced more negative/positive FODs during left/right rotation respectively and the size of negative FODs was larger (in terms of absolute value) during left than right rotation. Importantly, as opposed to adults, no significant differences in RNG between left and right head rotations were observed in children. Potential explanations for such age-related changes in the effect of active head rotation on RNG are discussed. Altogether, the present study confirms that numerical processing is spatially grounded in adults and suggests that its embodied aspect undergoes significant developmental changes.

Keywords: numerical cognition, embodied cognition, random number generation, active head rotation, developmental changes, children

INTRODUCTION

Knowledge and thinking are constrained by sensory-motor processes in that motor activities and other sensory-bodily experiences influence the cognitive processing of abstract concepts (Barsalou, 2008). The idea of such “embodied cognition” has become increasingly influential and numerical thinking can be considered as one principle example of it (Lakoff and Nunez, 2000).

According to the hierarchical model by Fischer and Brugger (2011; see also, Fischer, 2012), number processing is characterized by grounded, embodied and situated aspects. *Grounded numerical cognition* refers to the idea that numerical representations reflect the universal laws of the physical world in that small/large numbers are associated with lower/upper space respectively. This is supported by the observation that priming words linked to the lower (e.g., submarine) and upper (e.g., eagle) vertical space with small and large numbers respectively facilitated their treatment (Lachmair et al., 2014). *Embodied numerical cognition* is built on the basis of grounded cognition and suggests that number knowledge depends on spatial-directional learning experiences constituted by specific motor activities and other bodily sensory experiences. One example is the influence of finger counting habits on the association between numerical and spatial representations in healthy adults. While individuals who started counting on their left hand reliably associated small/large numbers with the left/right space respectively, no such effect was observed in right-starters (Fischer, 2008). Finally, *situated numerical cognition* suggests that number-space associations can be directly modulated by the current constraints and context of a situation, including both external stimuli as well as body posture. This level of knowledge representation is very flexible and instantly adapts to concurrent task demands. In that vein, Eerland et al. (2011) reported that participants’ numerical estimates were slightly smaller/larger when they were leaned toward the left/right respectively. Moreover, Loetscher et al. (2008) reported an effect of active motion on random number generation (RNG). Namely, participants produced more smaller/larger numbers while rotating their head toward the left/right respectively (see also Winter and Matlock, 2013).

The effect of movement on numerical production can be explained by spatial attention shifts along a hypothetical mental number line (MNL). According to the MNL hypothesis (for reviews see Dehaene, 1997; Hubbard et al., 2005), numbers are spatially represented along a horizontal axis with small/large numbers on its left/right respectively. The idea of the MNL was initially proposed following the observation of the *spatial-numerical association of response codes (SNARC) effect*, describing faster left-/right-sided responses for small/large digits respectively in binary classification tasks (Dehaene et al., 1993; Hoffmann et al., 2014; Georges et al., 2016). Motion-induced spatial attention shift on this so-called MNL would then bias the access to numerical magnitude representations, thereby explaining the effect of active head rotation on number selection (see Fischer and Shaki, 2014).

The robustness of the effect of motion on the reallocation of attention along the MNL was also further confirmed using bodily effectors other than the participant’s head. For instance, Loetscher et al. (2010) reported that the generation of smaller/larger numbers was preceded by left-/rightward eye movements respectively. Moreover, the selection of numbers during RNG depended on the direction of passive whole-body motion (Hartmann et al., 2012). In addition, Shaki and Fischer (2014) indicated that participants generated more small/large numbers when actively preparing to turn left-/rightward respectively. Interestingly, individuals were also more likely to turn to the left/right following the generation of a small/large number respectively. Cheng et al. (2015) also found that left-lateral arm turns facilitated the generation of smaller numbers relative to right-lateral turns. These findings thus collectively highlight the influence of motion on number processing in healthy adults, thereby providing evidence for the close link between numerical and spatial representations and the situatedness of their associations.

Nonetheless, despite the substantiation of situated numerical cognition in healthy adults, equally compelling evidence in children is sparse. To our knowledge, only Göbel et al. (2015) investigated the effect of spatially directional cues on RNG in 5- to 11-year old children. Concretely, they observed that lying on the left/right side of the body increased the generation of smaller/larger numbers respectively. It thus seems that directional cues can influence numerical production also in children, similarly to adults. However, it remains to be determined whether the generation of numbers at such earlier developmental stages can also be biased by active head rotation, as it has been repeatedly observed in healthy adults (Loetscher et al., 2008; Winter and Matlock, 2013; Cheng et al., 2015). Addressing this question should advance our understanding of spatial-numerical mappings in elementary school children and inform us on how their situatedness develops over the lifespan.

Aims

In the present study, we therefore aimed to determine the effect of active left/right head rotation on RNG not only in adults, but also in children. Children were recruited from 2nd, 3rd, and 4th grade of elementary school to be in line with the age range of the participants assessed in the study of Göbel et al. (2015), measuring the effect of spatially directional cues on RNG in 5- to 11-year-old children. This should enable us to replicate previous observations in adults and additionally inform us about whether the recently reported effect of static body position on RNG in children (Göbel et al., 2015) can be extended to active head rotation.

Finding evidence for an effect of active left/right head rotation on RNG not only in adults, but also in 7- to 11-year-old elementary school children would highlight potential similarities in spatial-numerical representations as well as in their situatedness across both age groups. In addition, it would suggest that the recently reported spatial bias in RNG observed in younger individuals (Göbel et al., 2015) is not specifically related to static body position. Conversely, the absence of an effect of active left/right head rotation on number processing in

children but not adults might indicate developmental changes in the spatial representation of numerical magnitudes. This would then be in line with studies indicating that estimation patterns on the number line task were fitted best by a logarithmic and linear function in children and adults respectively, suggesting an age-related log-to-linear shift in the representation of numerical magnitudes on the MNL (Booth and Siegler, 2006; Moeller et al., 2009). Alternatively, a potential null effect in children might suggest that these younger individuals do not yet activate spatial-numerical associations in tasks such as RNG, which do not involve any explicit magnitude judgments (e.g., van Galen and Reitsma, 2008). Furthermore, age-related differences in the effect of active head rotation on RNG could highlight potential developmental changes in the accessibility of the MNL. Children as opposed to adults might for instance not yet anchor number-space mappings onto an external reference frame when randomly generating numbers during head rotation (Crollen and Noël, 2015; Crollen et al., 2015; Nava et al., 2017). Finally, a potential null effect in the younger individuals could also simply be explained by the current paradigm of randomly generating numbers while actively rotating one's head. This dual-task scenario might compromise the working memory (WM) resources necessary for MNL activation, especially in children whose executive functions have not yet fully developed (Luciana and Nelson, 1998; De Luca et al., 2003; Best et al., 2009). This could then explain potential differences between the present outcomes and the previous findings by Göbel et al. (2015), who observed an effect of static body position on RNG already in elementary school children.

To quantify the effect of active head rotation on RNG, we computed (a) the mean of generated numbers and (b) the mean difference between each randomly generated number and its immediately preceding response, i.e., the first order difference (FOD). While the mean of generated numbers yields information about overall numerical selection preferences, the mean of FODs provides valuable insights into the way in which the generated numbers are selected on the MNL. More concretely, negative/positive FODs (reflecting descending/ascending steps in the generated numerical sequence) are indicative of left-/rightward spatial attention shifts along the MNL respectively, while smaller/larger FODs in terms of absolute value reflect narrow/wide spatial attention shifts along the MNL respectively regardless of direction.

The means of generated numbers and FODs are commonly used when studying spatial biases in RNG (e.g., Hartmann et al., 2012; Thompson et al., 2013; Winter and Matlock, 2013; Shaki and Fischer, 2014; Göbel et al., 2015). However, the sign (positive/negative) and absolute value (small/large) of FODs carry distinct information regarding spatial attention shifts along the MNL, namely their direction (left-/rightward) and size (narrow/wide) respectively. The relative contribution of these two factors to the overall mean of FODs consequently needs to be disentangled. Concretely, assessing whether e.g., a negative mean of FODs reflects (a) the generation of a higher total of descending steps (i.e., negative FODs) than ascending steps (i.e., positive FODs) and/or (b) the production of larger descending steps than ascending steps in terms of absolute value allows us

to investigate more thoroughly how active head rotation affects the reallocation of spatial attention along the MNL. In addition to reporting overall FOD values, we therefore assessed the effect of active left/right head rotation on the total of negative and positive FODs regardless of their numerical values as well as on the absolute value of negative and positive FODs.

MATERIALS AND METHODS

The experiment was reviewed and approved by the Ethics Review Panel of the University of Luxembourg. Adults signed a consent form and parental consent was obtained for the children prior to the start of the study

Participants

Children

In total, seventy children (36 female; mean age = 9.45 years; $SD = 1.10$; range = 7.8–11.9) were recruited from three Luxembourgish public elementary schools from the second ($N = 21$; age = 8.11; $SD = 0.35$), third ($N = 18$; age = 9.38; $SD = 0.60$) and fourth grade ($N = 31$; age = 10.4; $SD = 0.61$). None of the children had a history of learning disorders, such as dyslexia or dyscalculia. Data from the children reported in the present study were part of data collected in the framework of a bigger project including additional tasks not described hereafter.

Adults

Twenty-four participants (19 female; age = 23.3 years; $SD = 4.2$; range = 18–34) were recruited at the University of Luxembourg. They received a small compensation in exchange for their participation. None of them had a history of learning disorders, such as dyslexia or dyscalculia. They were all blind to the hypotheses of the experiment.

Procedure

Participants were asked to orally generate numbers between 1 and 30 as randomly as possible. To assist the participants in their understanding of “as randomly as possible,” we added the following sentence: “Imagine you have a bag in which there are thirty balls numbered 1–30 and whenever instructed you have to take a ball from the bag and tell me which number you see. After having said the number, you have to return the ball to the bag.” Subjects had to do this task while moving their head from left-to-right (i.e., right rotation) and from right-to-left (i.e., left rotation). They had their eyes covered with a mask during the entire task to prevent any distractions from their surroundings. The starting position of the head (head above left vs. right shoulder) was counterbalanced across participants. Participants were asked to generate the number halfway through their motion (i.e., when their head was aligned with their trunk), as opposed to when their head was fully turned toward the left/right side and as such had reached a static position (as in e.g., Loetscher et al., 2008). This was to clearly differentiate the current paradigm, investigating the effect of active head motion on RNG, from that of Göbel et al. (2015), who studied the effect of static left/right body position on numerical processing in children. The starting

of the head movement was announced through a beep given via a headset every 3.6 s. The average speed of the motion was therefore 0.14 Hz (i.e., one turn per 7.2 s). Rotational speed was slowed down compared to Loetscher et al. (2008) to provide the participants with sufficient time to generate numbers during active head rotation and to minimize the total of omissions and errors especially in the younger participants. The script of the generation task was running on Matlab on an 11-in. MacBook and the responses were recorded on a Maxxter Stereo Headset.

As previously done by Loetscher et al. (2008), 40 numbers had to be generated per condition (left, right), which resulted in 80 numbers in total. The session was divided into two blocks, thus resulting in 40 trials per block. To ensure that participants understood the task, a training session consisting of 16 trials preceded the actual experiment.

Data Analysis

First, we analyzed the total of *omissions* and *errors* during RNG. Responses were considered as erroneous if the generated number was outside the 1–30 range. We also quantified the overall randomness of number generation by computing the *redundancy score* (R score; Evans, 1978). The R score reflects the extent to which each response is generated with equal frequency. A score of 0% implies no redundancy, while a score of 100% indicates complete redundancy (i.e., all responses are identical). The latter calculation was achieved by a published computer program, freely downloadable at <http://www.lancs.ac.uk/staff/towse/rgcpage.html> (Towse and Neil, 1998). Assessing the effect of age group (adults vs. children) on these measures should inform us about potential age-related differences in overall task comprehension and performance. We also determined whether the total of omissions and errors during RNG depended on active left/right head rotation and/or its interaction with age group. It should be noted that left/right rotation refers to the left-/rightward motion during which the selected number had to be produced. Conversely, since measures of randomness in RNG, such as the R score, are not believed to rely on or directly index any numerical magnitude representations (Brugger, 1997), but supposedly predominantly depend on more general executive functions (Brugger, 1997; Baddeley, 1998; Peters et al., 2007; Terhune and Brugger, 2011), we did not assess the effect of rotation on the redundancy score.

To measure the effect of active head rotation on RNG, we referred to the study of Winter and Matlock (2013) and analyzed all correctly generated numbers as a continuous measure rather than binning them according to their magnitudes (i.e., smaller or larger than the mean of the number range; as in Loetscher et al., 2008). Two analyses were conducted based on this measure.

In a first step, we determined whether *the mean of correctly generated numbers* in each participant differed between active left/right head rotation.

In a second step, we focused on the arithmetic difference between each generated number and its immediately preceding response (i.e., the first order difference; FOD) and determined whether *the mean of FODs* in each participant differed between active left/right head rotation. In case a response was omitted or

outside the 1–30 range, the FODs between this incorrect/omitted response and its preceding as well as succeeding number were discarded from data analyses.

In general, FODs can be classified depending on two factors: (1) their sign (positive vs. negative) and (2) their absolute numerical value (small vs. large). In spatial terms, the sign of the FOD reflects the direction of the step on the MNL. While positive FODs, indexing an ascending step in the generated numerical sequence, correspond to a “rightward” shift along the MNL, negative FODs, reflecting a descending step in the generated numerical sequence, correspond to a “leftward” shift on the MNL. Conversely, the absolute numerical value of the FOD reflects the size of the step on the MNL regardless of its ascending or descending direction. The overall mean of FODs thus depends on the interplay between these two factors.

To disentangle the relative contribution of these two factors to the mean of FODs, we performed two additional analyses.

Firstly, we determined whether *the total of FODs* differed depending on their positive/negative sign – henceforth referred to as “direction” as it reflects the ascending/descending direction of the step in the generated numerical sequence – during left/right rotation. In other terms, we assessed the effects of direction and rotation on the total of FODs. More concretely, we compared the total of positive and negative FODs during both left and right rotations regardless of their numerical value. We hypothesized that positive FODs should outnumber negative FODs during right rotation, but vice-versa during left rotation. Moreover, participants should generate more positive/negative FODs during right/left than left/right rotations respectively. It should be noted that the total of FODs is a continuous variable ranging from 0 to 39 during both left and right rotations (i.e., the total of 40 numbers generated per left/right rotation minus one). In addition, it is worth mentioning that the totals of positive and negative FODs should in theory be inversely proportional. More concretely, more positive FODs should be associated with less negative FODs such that the total of FODs always adds up to 39. Nonetheless, this was practically not the case in the present investigation considering the exclusion of FODs preceding as well as succeeding erroneous and omitted responses. Moreover, FODs of zero, resulting from the repetition of the same number on two (or more) consecutive trials, could not be considered for the current analysis. The totals of positive and negative FODs thus ranged from 24 to 39 and 27 to 39 during left and right rotation respectively, entailing that negative and positive FODs were practically not directly inversely proportional in the present study. As such, it is important to include “direction” as an additional factor in the ANOVA rather than simply assessing only the effect of rotation on either positive or negative FODs.

Secondly, we ascertained whether direction and/or rotation affected *the mean absolute value of FODs*. More concretely, we compared the means of positive and negative FODs in terms of absolute value during both left and right rotations. We hypothesized that positive FODs should be larger in terms of absolute value than negative FODs during right rotation, but vice-versa during left rotation. Moreover, participants should perform larger positive/negative FODs (in terms of absolute value) during right/left than left/right rotations.

To measure potential age-related changes in the effect of active head rotation on RNG, the above analyses were conducted including age group (adults vs. children) as between-subject factor. In case of a significant interaction between age group and one of the within-subject variables (i.e., rotation and direction), two separate ANOVAs were subsequently performed – one for each age group. When only focusing on the subgroup of children, grade was additionally added as a between-subject factor in all the analyses. This was mainly to exclude the possibility that any potential interaction effects between age group and the within-subject variables rotation and/or direction on the different dependent variables were driven only by a certain grade.

Considering that individuals' counting strategies might be affected by their initial starting position (see Towse and Cheshire, 2007), each of the following analyses was initially conducted including starting orientation as an additional between-subject variable. Since starting orientation did, however, not have any main or interaction effects, we decided to drop this variable from data analysis. All analyses reported below were thus conducted without starting orientation as between-subject factor.

An alpha of 0.05 was used as the cut-off for significance (i.e., the null hypothesis was rejected if $p < 0.05$) in all the following analyses.

RESULTS

All descriptives can be found in Table 1.

Preliminary Analyses

The Total of Omissions as a Function of Rotation in Children and Adults

A 2×2 mixed ANOVA on the total of omissions including rotation and age group as within- and between-subject factors respectively indicated a main effect of age group [$F(1,92) = 9.97$; $p = 0.002$; $\eta_p^2 = 0.1$], with children omitting responses on significantly more trials than adults (children: $\bar{x} = 1.9$; $SD = 2.8$ vs. adults: $\bar{x} = 0.08$; $SD = 0.28$). The total of omissions did, however, not differ between left and right rotation [$F(1,92) = 0.01$; $p = 0.94$; $\eta_p^2 = 0.00$] and there was no interaction between rotation and age group [$F(1,92) = 0.01$; $p = 0.94$; $\eta_p^2 = 0.00$]. In children, a 2×3 mixed ANOVA including rotation and grade as within- and between-subject factors respectively revealed no effect of grade [$F(2,67) = 0.77$; $p = 0.47$; $\eta_p^2 = 0.02$] and there was no interaction between grade and rotation [$F(2,67) = 0.78$; $p = 0.46$; $\eta_p^2 = 0.02$].

The Total of Errors as a Function of Rotation in Children and Adults

A 2×2 mixed ANOVA on the total of errors including rotation and age group as within- and between-subject factors respectively indicated a main effect of age group [$F(1,92) = 7.48$; $p = 0.01$; $\eta_p^2 = 0.08$], with children generating more numbers outside the 1–30 range than adults (children: $\bar{x} = 1.4$; $SD = 2.5$ vs. adults: $\bar{x} = 0.0$; $SD = 0.0$). However, there was no main effect of rotation [$F(1,92) = 0.95$; $p = 0.33$; $\eta_p^2 = 0.01$] and also no interaction between rotation and age group [$F(1,92) = 0.95$; $p = 0.33$; $\eta_p^2 = 0.01$]. A 2×3 mixed ANOVA including rotation and grade

as within- and between-subject factors respectively indicated that grade significantly affected the total of errors [$F(2,67) = 4.00$; $p = 0.02$; $\eta_p^2 = 0.11$]. *Post hoc* pairwise comparisons revealed that 2nd graders generated more numbers outside the specified range than 4th graders [2nd grade: $\bar{x} = 2.43$ vs. 4th grade: $\bar{x} = 0.55$; $t(24.46) = 2.57$; $p = 0.02$; Cohen's $d = 0.77$]. There was, however, no interaction between grade and rotation [$F(2,67) = 0.05$; $p = 0.95$; $\eta_p^2 = 0.001$].

Only correct responses within the 1–30 range (95.9% in children vs. 99.9% in adults) were considered for all subsequent analyses.

The Randomness Quality in Children and Adults

The redundancy score across all participants was 9.84 ($SD = 6.05$). A one-way ANOVA including age group as between-subject factor revealed a main effect [$F(1,92) = 4.94$; $p = 0.03$; $\eta_p^2 = 0.05$], with adults generating more random numerical sequences than children (adults: R score = 7.52; $SD = 4.20$ vs. children: R score = 10.64; $SD = 6.40$). In children, a one-way ANOVA indicated that the R score did, however, not differ depending on grade [$F(2,67) = 2.75$; $p = 0.07$; $\eta_p^2 = 0.08$].

The Mean of Generated Numbers as a Function of Rotation in Children and Adults

A 2×2 mixed ANOVA on the mean of correctly generated numbers including rotation and age group as within- and between-subject factors respectively did not reveal a main effect of rotation [$F(1,92) = 3.67$; $p = 0.06$; $\eta_p^2 = 0.04$] or age group [$F(1,92) = 0.06$; $p = 0.81$; $\eta_p^2 = 0.001$]. Nonetheless, a significant interaction between rotation and age group was observed [$F(1,92) = 5.08$; $p = 0.03$; $\eta_p^2 = 0.05$]. A follow-up repeated measures ANOVA in adults revealed that the effect of rotation was significant [$F(1,23) = 5.21$; $p = 0.03$; $\eta_p^2 = 0.19$; see Figure 1]. Namely, adults generated on average smaller numbers during left ($\bar{x} = 13.63$; $SD = 1.92$) than right rotation ($\bar{x} = 14.34$; $SD = 1.80$). Conversely, in children, a follow-up 2×3 mixed ANOVA including rotation and grade as within- and between-subject factors respectively did not indicate an effect of rotation [$F(1,67) = 0.003$; $p = 0.96$; $\eta_p^2 = 0.00$; see Figure 1], suggesting that the mean of generated numbers did not significantly differ depending on left ($\bar{x} = 14.14$; $SD = 2.46$) or right rotation ($\bar{x} = 14.08$; $SD = 2.46$). In the latter participants, there was also no effect of grade [$F(2,67) = 1.31$; $p = 0.28$; $\eta_p^2 = 0.04$] nor did the interaction between rotation and grade reach significance [$F(2,67) = 0.76$; $p = 0.47$; $\eta_p^2 = 0.02$].

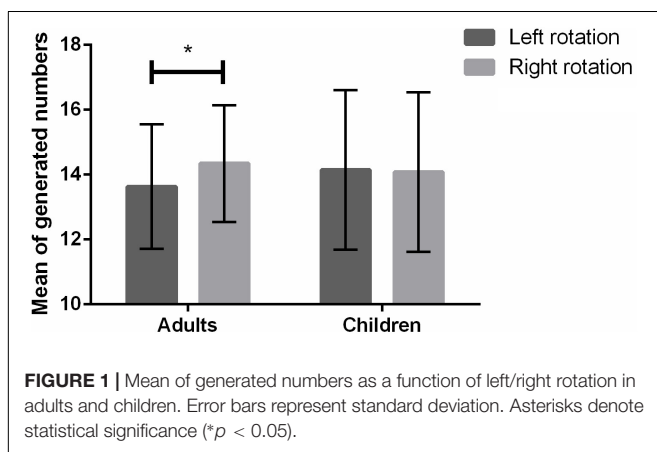
The Mean of First Order Differences as a Function of Rotation in Children and Adults

A 2×2 mixed ANOVA on the mean of the differences between each generated number and its immediately preceding response (i.e., the FOD) including rotation and age group as within- and between-subject factors respectively did not indicate a main effect of age group [$F(1,92) = 2.67$; $p = 0.11$; $\eta_p^2 = 0.03$]. However, a main effect of rotation [$F(1,92) = 5.26$; $p = 0.02$; $\eta_p^2 = 0.05$] as

TABLE 1 | Descriptive information.

Variables	Children		Adults	
	Left rotation	Right rotation	Left rotation	Right rotation
Total of omissions	0.94 (1.52)	0.96 (1.44)	0.04 (0.20)	0.04 (0.20)
Total of errors	0.6 (1.17)	0.8 (1.50)	0	0
R score	10.64 (6.40)		7.52 (4.20)	
Mean of generated numbers	14.14 (2.46)	14.08 (2.46)	13.63 (1.92)	14.34 (1.80)
Mean of FODs	0.27 (1.49)	0.23 (1.45)	−0.69 (1.62)	0.93 (1.50)
Total of negative FODs	14.74 (4.61)	14.44 (4.09)	17.71 (4.62)	14.83 (3.84)
Total of positive FODs	21.39 (5.08)	21.84 (4.03)	21.13 (4.47)	24.00 (3.74)
Mean absolute value of negative FODs	7.14 (2.44)	8.13 (3.03)	9.82 (3.02)	8.59 (2.96)
Mean absolute value of positive FODs	5.53 (1.94)	5.66 (2.11)	6.70 (2.13)	6.98 (1.81)

Standard deviations are shown in parentheses. The R score was only compared between the different age groups, but not between left-/right-sided rotation.



well as a significant interaction between rotation and age group [$F(1,92) = 5.88$; $p = 0.02$; $\eta_p^2 = 0.06$] were revealed. A follow-up repeated measures ANOVA in adults indicated that the effect of rotation was significant [$F(1,23) = 6.59$; $p = 0.02$; $\eta_p^2 = 0.22$; see **Figure 2A**]. Namely, the mean of FODs was significantly smaller during left ($\bar{x} = -0.69$) than right rotation ($\bar{x} = 0.93$). Conversely, in children, a follow-up 2×3 mixed ANOVA including rotation and grade as within- and between-subject factors respectively revealed no main effect of rotation [$F(1,67) = 0.008$; $p = 0.93$; $\eta_p^2 = 0.00$; see **Figure 2A**], indicating no significant differences in the mean of FODs depending on left ($\bar{x} = 0.27$) or right rotation ($\bar{x} = 0.23$). Moreover, there was no main effect of grade [$F(2,67) = 0.15$; $p = 0.86$; $\eta_p^2 = 0.01$] and no interaction between grade and rotation on the mean of FODs in children [$F(2,67) = 0.47$; $p = 0.62$; $\eta_p^2 = 0.01$].

The Total of First Order Differences as a Function of Direction and Rotation in Children and Adults

A $2 \times 2 \times 2$ mixed ANOVA on the total of FODs including age group as between-subject factor and direction (ascending/positive vs. descending/negative) as well as rotation as within-subject variables did not indicate a main effect of rotation [$F(1,92) = 0.43$; $p = 0.52$; $\eta_p^2 = 0.01$], but a significant effect of direction was revealed [$F(1,92) = 66.47$; $p < 0.001$;

$\eta_p^2 = 0.42$]. Namely, a higher total of positive than negative FODs was observed across all participants, indicating that individuals generated more ascending than descending steps regardless of age. Moreover, a main effect of age group was revealed [$F(1,92) = 21.29$; $p < 0.001$; $\eta_p^2 = 0.19$], with the total of FODs being significantly higher in adults than children. This confirms the higher total of omissions and errors in the latter participants (see above).

Most interestingly, however, a significant interaction between direction and rotation was observed [$F(1,92) = 7.95$; $p = 0.006$; $\eta_p^2 = 0.08$], which additionally depended on age group [$F(1,92) = 4.68$; $p = 0.033$; $\eta_p^2 = 0.048$]. A follow-up 2×2 repeated measures ANOVA in adults including direction and rotation as within-subject factors indicated a significant interaction between these variables [$F(1,23) = 5.46$; $p = 0.03$; $\eta_p^2 = 0.19$; see **Figure 2B**]. More concretely, the total of negative FODs was higher during left ($\bar{x} = 17.71$) than right rotation [$\bar{x} = 14.83$; $F(1,23) = 5.44$; $p = 0.03$; $\eta_p^2 = 0.19$], while positive FODs were more numerous during right ($\bar{x} = 24.00$) than left rotation [$\bar{x} = 21.13$; $F(1,23) = 5.47$; $p = 0.03$; $\eta_p^2 = 0.19$]. In addition, positive FODs significantly out-numbered negative FODs during right [positive: $\bar{x} = 24.00$; negative: $\bar{x} = 14.83$; $F(1,23) = 35.16$; $p < 0.001$; $\eta_p^2 = 0.61$] but not left rotation [positive: $\bar{x} = 21.13$; negative: $\bar{x} = 17.71$; $F(1,23) = 3.40$; $p = 0.08$; $\eta_p^2 = 0.13$]. As opposed to adults, a follow-up $2 \times 2 \times 3$ mixed ANOVA in children including rotation and direction as within-subject factors and grade as between-subject variable did not indicate an interaction between direction and rotation [$F(1,67) = 0.69$; $p = 0.41$; $\eta_p^2 = 0.01$; see **Figure 2C**]. Positive FODs were more numerous than negative FODs regardless of left [positive: $\bar{x} = 21.39$; negative: $\bar{x} = 14.74$; $F(1,67) = 32.09$; $p < 0.001$; $\eta_p^2 = 0.32$] or right rotation [positive: $\bar{x} = 21.84$; negative: $\bar{x} = 14.44$; $F(1,67) = 60.23$; $p < 0.001$; $\eta_p^2 = 0.47$]. Moreover, no significant differences between left and right rotations were observed for the totals of negative [$F(1,67) = 0.45$; $p = 0.50$; $\eta_p^2 = 0.01$] or positive FODs [$F(1,67) = 0.95$; $p = 0.34$; $\eta_p^2 = 0.01$]. The absence of an interaction between direction and rotation in children did not depend on grade [$F(2,67) = 0.22$; $p = 0.81$; $\eta_p^2 = 0.01$] and there was no main effect of grade on the total of FODs [$F(2,67) = 1.55$; $p = 0.22$; $\eta_p^2 = 0.04$].

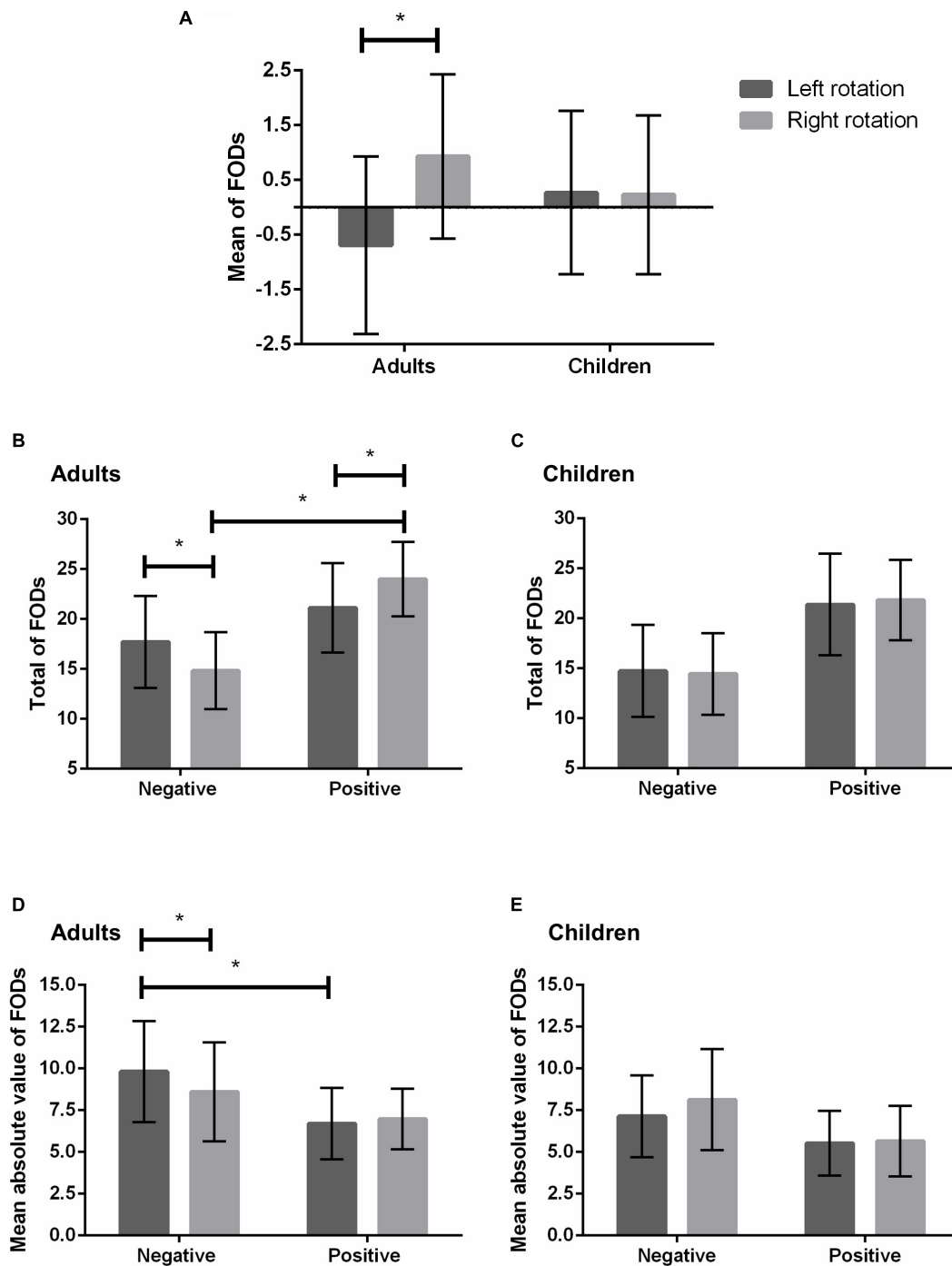


FIGURE 2 | First order differences as a function of left/right rotation in adults and children. The mean of FODs (A). The total of negative and positive FODs in adults (B) and children (C). The mean absolute value of negative and positive FODs in adults (D) and children (E). Error bars represent standard deviation. Asterisks denote statistical significance (* $p < 0.05$).

The Mean Absolute Value of First Order Differences as a Function of Direction and Rotation in Children and Adults

A $2 \times 2 \times 2$ mixed ANOVA on the mean absolute value of FODs including age group as between-subject factor and direction as

well as rotation as within-subject variables indicated no main effect of rotation [$F(1,92) = 0.05$; $p = 0.82$; $\eta_p^2 = 0.001$], but a significant effect of direction [$F(1,92) = 59.87$; $p < 0.001$; $\eta_p^2 = 0.39$]. In general, participants generated larger negative ($\bar{x} = 8.15$; $SD = 2.57$) than positive FODs ($\bar{x} = 5.94$; $SD = 1.88$)

in terms of absolute value (i.e., individuals generally performed larger descending steps). Moreover, a main effect of age group was observed [$F(1,92) = 11.35$; $p = 0.001$; $\eta_p^2 = 0.11$] in that larger FODs were performed by adults ($\bar{x} = 7.74$; $SD = 1.94$) than children ($\bar{x} = 6.22$; $SD = 1.75$).

Most importantly, however, we found a significant interaction between direction, rotation and age group [$F(1,92) = 7.57$; $p = 0.007$; $\eta_p^2 = 0.08$]. A follow-up 2×2 repeated measures ANOVAs in adults including direction and rotation as within-subject factors indicated that the interaction between direction and rotation was significant [$F(1,23) = 4.59$; $p = 0.04$; $\eta_p^2 = 0.17$; see **Figure 2D**]. Namely, the main effect of direction with larger negative than positive FODs was more pronounced during left [negative $\bar{x} = 9.82$; positive $\bar{x} = 6.70$; $F(1,23) = 31.65$; $p < .001$; $\eta_p^2 = 0.58$] than right rotation [negative $\bar{x} = 8.59$; positive $\bar{x} = 6.98$; $F(1,23) = 7.29$; $p = 0.013$; $\eta_p^2 = 0.24$]. Moreover, a main effect of rotation was observed for negative FODs [$F(1,23) = 4.66$; $p = 0.04$; $\eta_p^2 = 0.17$] in that the latter were significantly larger in terms of absolute value during left ($\bar{x} = 9.82$) than right rotation ($\bar{x} = 8.59$). In children, a follow-up $2 \times 2 \times 3$ mixed ANOVA including rotation and direction as within-subject factors and grade as between-subject variable indicated a main effect of rotation [$F(1,67) = 7.56$; $p = 0.008$; $\eta_p^2 = 0.10$], with children generating larger FODs when turning their heads right- ($\bar{x} = 6.4$) compared to leftward ($\bar{x} = 6.04$). As opposed to adults, the effect of rotation did, however, not depend on direction [$F(1,67) = 2.3$; $p = 0.13$; $\eta_p^2 = 0.03$; see **Figure 2E**]. The absence of an interaction between rotation and direction in children was not affected by grade [$F(2,67) = 1.77$; $p = 0.18$; $\eta_p^2 = 0.05$] and there was no overall effect of grade [$F(2,67) = 2.56$; $p = 0.08$; $\eta_p^2 = 0.07$].

DISCUSSION

The present study aimed to determine whether active left/right head rotation biases RNG not only in adults, but also in 7- to 11-year-old elementary school children. This should inform us about whether the recently reported effect of static body position on RNG in children (Göbel et al., 2015) can be extended to active head rotation. Overall, this will further advance our understanding of spatial-numerical mappings in elementary school children and the lifespan development of their situatedness.

In line with previous findings (Loetscher et al., 2008; Winter and Matlock, 2013; Cheng et al., 2015), adults produced on average smaller numbers during left than right rotation. In addition, the mean of FODs was smaller when rotating the head left- as opposed to rightward. Considering that the average FOD potentially depends not only on the total of descending and ascending steps in the generated numerical sequence, but also on their respective absolute values, we additionally studied the effects of rotation on the total as well as the absolute value of negative and positive FODs. This provides further information on how active head rotation affects spatial attention shifts along the MNL. Interestingly, the smaller mean of FODs during left than right rotation reflected the generation of significantly larger descending than

ascending steps in terms of absolute value, while the larger mean of FODs during right than left rotation was mainly due to the production of a higher total of ascending than descending steps. When looking at it from a different angle, participants produced more descending steps during left than right rotation, while ascending steps were more numerous when moving the head right- as opposed to leftward. This suggests that participants shifted their attentional focus more often toward the left/right along the MNL when rotating their heads in the left/right direction respectively. In addition, the size of descending steps was larger (in terms of absolute value) during left than right movement. Overall, these findings highlight the close link between numerical and spatial representations, likely encoded in overlapping brain circuits in the posterior parietal cortex and particularly in areas in and around the intraparietal sulcus (for reviews, see Hubbard et al., 2005, 2009; for the “neuronal recycling” hypothesis, see Dehaene, 2005; Dehaene and Cohen, 2007). Moreover, the present findings provide further evidence for the situatedness of spatial-numerical interactions in adults.

Importantly, as opposed to adults, we did not observe a significant influence of active head rotation on RNG in children. The absence of a significant effect in the latter participants did also not depend on grade. Although the absence of evidence for a significant difference in RNG between left and right rotation in children should not be directly considered as evidence of absence of an effect of active head rotation on RNG in the younger participants, the present findings suggest that the spatial bias in RNG during active head motion observed in adults likely only emerges at later developmental stages, at the earliest after 4th grade. In general, the observed null effect in 2nd to 4th graders might have several reasons, which will be discussed in the following paragraphs.

First, the absence of a significant effect of active left/right head rotation on number processing in children might indicate that these younger individuals do not yet represent numerical magnitudes in a spatial format akin to a MNL, as it is likely the case in adults. This assumption is supported by the observation that reading direction affected the orientation of spatial-numerical mappings on the MNL (Shaki et al., 2009), suggesting that number-space associations only gradually arise after formal schooling through reading acquisition (see also, Berch et al., 1999; Zebian, 2005; White et al., 2012). In line with this view, Ninaus et al. (2017) recently observed an age-related increase in the SNARC effect. These findings thus collectively suggest that number-space associations probably only arise later in life through embodied spatially directional experiences such as reading and writing direction.

Nonetheless, the idea that spatial-numerical interactions only arise after formal schooling through reading acquisition was refuted by studies evidencing number-space associations also in preliterate children. Namely, Hoffmann et al. (2013) reported a SNARC effect in a color judgment task already in 5.5-year-old preschoolers. Patro and Haman (2012) even observed a SNARC-like effect in 4-year-old children in that they associated small/large non-symbolic numerosities with the left/right respectively. In addition, most preschoolers

already add, subtract and count from left-to-right (Opfer et al., 2010; Opfer and Furlong, 2011; Shaki et al., 2012). Interestingly, left-to-right counting was only observed in children growing up in England, while Palestinian preschoolers mainly counted from right-to-left (Shaki et al., 2012). Number-space associations thus likely emerge much earlier in life through directionally relevant cultural experiences. Interestingly, some studies even reported number-space associations in infants and neonates (de Hevia et al., 2006, 2014a,b; de Hevia and Spelke, 2009, 2010; Lourenco and Longo, 2010), thereby suggesting their innateness. The null effect of active left/right head rotation on number processing in children is thus not likely to be explained by children's lack of spatial-numerical interactions.

A more likely explanation for the observed discrepancy between adults and children could be developmental changes in the spatial representation of numerical magnitudes. Interestingly, estimation patterns on the number line task were fitted best by a logarithmic and linear function in children and adults respectively, suggesting an age-related log-to-linear shift in the representation of numerical magnitudes on the MNL (Booth and Siegler, 2006; Moeller et al., 2009). Within a logarithmic representation, small numbers are spaced further apart than larger ones (Simms et al., 2016). Children should thus have better access than adults to relatively smaller numerical magnitudes, given their extended representations on the MNL. We did, however, not observe a main effect of age group on the mean of generated numbers, suggesting no age differences in the selection of smaller numerical magnitudes and as such spatial-numerical representations in the current sample. Moreover, previous studies indicated that performances on the 0-to-100 number line estimation task can already be best explained by a linear model from 2nd grade onward (Siegler and Opfer, 2003; Siegler and Booth, 2004). Children in the present study, especially those attending 3rd and 4th grade, thus probably featured mostly linear spatial-numerical representations. Finally, it is also worth noting that although estimation patterns in the number line task are usually interpreted as an indication of the logarithmic or linear nature of numerical magnitude representations (e.g., Siegler and Opfer, 2003; Laski and Siegler, 2007; Opfer and Siegler, 2007), performances on this task might not directly index scaling of the MNL representation in an isomorphic way. Number line estimation performances might more likely index number knowledge (Ebersbach et al., 2008), understanding of the place-value structure (e.g., Moeller et al., 2009), the adoption of certain solution strategies (Barth and Paladino, 2011; Cohen and Blanc-Goldhammer, 2011; Slusser et al., 2013) or attention processes (Anobile et al., 2012). Consequently, rather than reflecting a developmental change in the underlying spatial-numerical representations, the age-related log-to-linear shift in the fit of number line estimation performances might indicate the adoption of different resolution strategies in children and adults. It is therefore unclear whether children and adults feature different spatial-numerical representations. Developmental changes in the latter might thus not be the reason underlying age-related differences in the effect

of active head rotation on RNG. Moreover, if this were the case, the effect of head motion on RNG in children should have depended on grade, with 3rd and 4th graders showing similar spatial biases in RNG during rotation than adults, given their already mostly linear numerical magnitude representations.

The null effect in children, as opposed to adults, could, however, potentially be explained by age-related changes in the activation of number-space mappings on the MNL. Children might simply not yet activate spatial-numerical associations in tasks such as RNG, which do not involve any explicit magnitude judgments. This idea is in line with results from van Galen and Reitsma (2008), who observed that younger children only displayed a SNARC effect during explicit magnitude classifications, but not when numerical magnitude information was task-irrelevant during parity judgments. Nonetheless, as already mentioned before, Hoffmann et al. (2013) reported a SNARC effect in a numerical magnitude-irrelevant color judgment task even in preschoolers at the age of 5.5 years. Moreover, Chinese children were shown to display a parity SNARC effect already in Kindergarten at the age of 5.8 years (Yang et al., 2014). The latter findings thus suggest that children activate spatial-numerical representations on the MNL even when numerical magnitude information is not directly task-relevant. As such, inefficient activation of the MNL during RNG in children might not account for the absence of a significant effect of active left/right head rotation on number production in the latter individuals. It should, however, be noted that Göbel et al. (2015) failed to observe a relation between the parity SNARC effect and the spatial bias in RNG in adults, suggesting that these effects might arise from different underlying spatial-numerical representations. As such, evidence for number-space mappings during numerical magnitude-irrelevant parity judgments might not necessarily suggest the activation of spatial-numerical representations also during RNG. Consequently, it cannot be refuted that children, as opposed to adults, did not activate numerical representations on the MNL while randomly selecting numbers in the present study, which could then explain the null effect of active left/right head rotation.

Another possible explanation might be that the activation pattern of spatial-numerical representations does not yet depend on situated factors at earlier developmental stages. Nonetheless, spatially directional cues such as left/right body position were previously shown to increase the generation of smaller/larger numbers respectively already in 5 to 11-year-old children (Göbel et al., 2015). Moreover, simply observing left-to-right or right-to-left reading from storybooks instantaneously affected the counting direction of 3 to 5-year old preliterate children in line with the direction of observed reading (Göbel et al., 2017). The activation of spatial-numerical representations on the MNL thus seems to be flexibly modulated by situational demands also in children. The lack of situatedness of spatial-numerical associations in children therefore unlikely explains the current findings.

Children might, however, access their number-space mappings in a different way than adults. Developmental

changes in the accessibility of the MNL could then explain age-related differences in the effect of active head rotation on RNG. In this vein, Towse et al. (2014) reported that children featured different number preferences than adults during RNG. Namely, while adults showed a reliable and systematic bias toward the selection of smaller numbers, 8- to 11-year-old children preferentially generated larger numbers. The authors also evidenced a relation between age and the strength of the small number bias, suggesting a developmental increase in the preference for the selection of smaller numerical magnitudes. Adults were also shown to generate both ascending and descending numerical sequences, while children tended to produce mostly ascending sequences (Towse et al., 2014). Reluctance toward the generation of descending steps in children might not only explain their greater preferences for the selection of larger numbers in the study of Towse et al. (2014), but also potentially account for the absence of a significant effect of active head rotation on RNG in the present investigation.

Moreover, children likely anchor number-space mappings onto different spatial reference frames than adults. Namely, 6-year-old children did not display a SNARC effect when their hands were crossed (Nava et al., 2017), while sighted adults featured regular number-space associations regardless of hand posture (Dehaene et al., 1993; Crollen and Noël, 2015; Crollen et al., 2015). These findings suggest that younger, as opposed to older, individuals do not yet exclusively rely on an external object-centered reference frame when spatially representing numbers. They might rather depend on both internal body-centered and external frames of reference for mapping numbers onto space. The anchoring of spatial-numerical representations solely onto external coordinates might thus only gradually arise with increasing age.

Interestingly, number-space associations in 6-year-olds, but not adults, also depended on visual feedback in that no SNARC effect was observed when children were blindfolded (Nava et al., 2017). The ability to anchor numerical concepts onto an external spatial reference frame thus seems to depend on the availability of visual cues, especially at earlier developmental stages. The importance of visual experience for the development of an adult-like anchoring of numerical representations onto external space is also in line with findings in early blind adults. Namely, these individuals showed a reversed SNARC effect with crossed hands, indicating the adoption of a hand-centered reference frame during number processing (Crollen et al., 2013). Regarding these findings, children in the present study might not have been able to anchor number-space mappings onto an external reference frame when randomly generating numbers during head rotation, especially since they were blindfolded. The lack of visual feedback either completely kept them from accessing their spatial-numerical representations or induced them to rely on a rather head-centered frame of reference. This, in turn, might have masked the effect of active head rotation on RNG in the latter population. Conversely, adults probably used external spatial coordinates in that they coded numbers spatially with respect to their head facing straightforward. Left-/rightward head turns away from this position might then

have induced associated spatial attention shifts on the MNL, leading to the generation of smaller/larger numbers during left/right rotation respectively. It should, however, be noted that the spatial bias in RNG in the study of Göbel et al. (2015) was evidenced despite the children having their eyes closed. This thus suggests that these younger individuals were probably able to rely on an external reference frame even in the absence of visual input. The potential reliance on a body-centered spatial reference frame during RNG due to the absence of visual feedback at earlier developmental stages is thus unlikely to account for the present null effect in the younger individuals.

An alternative explanation for the null effect in children might be that although these younger individuals could use external spatial coordinates, similarly to adults, the current instruction to generate the number while facing straightforward directed their spatial attention toward where their head was positioned at the time of number generation (i.e., straight ahead). Consequently, their left/right head turns might not have been associated with respective spatial attention shifts on the MNL. This, in turn, could then explain the absence of a significant difference in RNG between left and right rotation in the younger individuals. This explanation could also account for the spatial biases in RNG observed in the study of Göbel et al. (2015), considering that the children were positioned on their left/right and thus facing in the corresponding direction. Nonetheless, the hypothesis that spatial attention was focused straight ahead due to task instructions would anticipate a null effect also in adults, since both adults and children received the same instructions in the present investigation. It could, however, still be that the spatial attention of older as opposed to younger individuals was not restricted toward where their head was positioned at the time of number selection.

Another reason for the discrepancy between the present findings in children and those of Göbel et al. (2015) could lie in the way the effects of space were assessed. While Göbel et al. (2015) determined the impact of static left/right body orientation on RNG, we assessed the effect of active left/right head motion. In addition, it needs to be reminded that in the current set-up participants had to generate a random number during motion, while classically in the literature numbers are produced once the movement has finished (see e.g., Loetscher et al., 2008). Since participants had to generate numbers while simultaneously moving their heads left-/rightward, the current paradigm can be considered as a dual-task and was therefore probably more difficult than that implemented in previous studies. Randomly generating numbers in a situation involving lateral head turns as well as the fact that the numbers had to be produced during motion (as opposed to when the head had reached a static left/right position) might have placed additional demands on the WM system, already strained by the RNG task in itself (Jahanshahi et al., 1998; Hamdan et al., 2004). Considering that WM and executive functions have not yet fully developed in children (Luciana and Nelson, 1998; De Luca et al., 2003; Best et al., 2009), the latter participants might have been particularly negatively affected by this dual-task situation. This interpretation is supported

by the greater number of omissions in children compared to adults. In addition, children featured a higher redundancy score than adults, indicating that they selected numbers less randomly. Considering that this measure is interpreted to rely on general executive functions, such as the ability to suppress response preferences created by one's own previous output (Brugger, 1997; Baddeley, 1998; Peters et al., 2007; Terhune and Brugger, 2011), this further endorses the assumption that executive processing was particularly strained in children. Since number-space associations were previously shown to depend on available WM resources in that no SNARC effect was observed under increased WM load (Herrera et al., 2008; van Dijck et al., 2009), compromised WM resources especially in children might have prevented them from accessing spatial-numerical representations during RNG and as such precluded any spatial bias in their numerical magnitude selection during active head rotation. This could then account for the null effect in the present study, even though spatial biases were previously evidenced by Göbel et al. (2015). Overall, this interpretation further strengthens the important role of WM in the association between spatial and numerical concepts (Herrera et al., 2008; van Dijck et al., 2009, 2014; van Dijck and Fias, 2011; Ginsburg et al., 2014; Abrahamse et al., 2016; Fias and van Dijck, 2016). Considering that WM ability considerably increases between adolescence and adulthood, especially for tasks requiring retention during distraction (Fry and Hale, 2000; Gathercole et al., 2004; Ullman et al., 2014), the spatial bias in RNG during active head rotation might only arise in older children attending high-school. This would then also account for the fact that school grade did not influence the effect of active head rotation on RNG in the present group of elementary school children. In other terms, it would provide an explanation for why number selection did not significantly differ between active left/right head rotation, even in the oldest children of the current sample.

Future Studies

To verify whether the null effect in children might be explained by their logarithmic as opposed to linear numerical magnitude representations, future studies could additionally administer a number line estimation task assessing the linearity of numerical magnitude representations. Accordingly, RNG should be least affected by active head rotation in those children featuring more logarithmic representations. The latter children should also generally produce more smaller numbers compared to their age-matched peers.

An interesting idea might also be to prime the activation of the MNL by instructing children to imagine numbers on a ruler while performing the RNG task (see Loetscher et al., 2008). This should yield valuable information regarding whether the absence of a significant effect of active head rotation on RNG in children might be explained by inefficient activation of spatial-numerical representations on the MNL during task completion.

Future studies might also envisage to replicate the present investigation without blindfolding participants. This should unravel whether the absence of visual feedback

and the anchoring of numerical magnitudes onto head-centered as opposed to extra-corporal spatial coordinates in children could have accounted for the absence of a significant effect of active head rotation on RNG in the latter individuals.

Finally, to determine whether the dual-task situation and the associated compromise in available WM resources contributed to the null effect in children, one could additionally assess the children's WM capacity. Accordingly, a null effect might only be observed in those children with weaker WM performances, while active left/right head rotation might lead to the generation of smaller/larger numbers respectively in those children with higher WM capacity, similarly to adults. Alternatively, one could assess RNG performances in a static experimental set-up not involving any left/right head motion. Considering that higher executive functions as well as WM are associated with better randomness quality (Brugger, 1997; Baddeley, 1998; Peters et al., 2007), finding evidence for better RNG performances in terms of the R score as well as the total of errors and omissions in the absence of active head rotation could then substantiate the hypothesis that WM resources were indeed likely reduced in the current dual-task paradigm, which in turn might have potentially accounted for the absence of a significant difference in RNG between left and right rotation in children.

CONCLUSION

To conclude, we replicated previous findings showing an effect of active head rotation on the randomization of numbers in adults. Adults generated on average smaller numbers and the mean of FODs was smaller during left than right rotation. Importantly, by additionally studying the effects of rotation on the total as well as the absolute value of negative and positive FODs, the present study significantly advanced our understanding of how spatially directional cues such as active head rotation affect step generation and as such spatial attention shifts along the MNL. Participants produced more descending/ascending steps during left/right head rotation respectively, indicating that they shifted their attentional focus more often toward the left/right along the MNL when rotating their heads in the corresponding direction. In addition, the size of descending steps was larger (in terms of absolute value) during left than right rotation. As opposed to adults, RNG in elementary school children did not significantly differ between active left/right head rotation. Future studies should determine whether such age-related differences can be explained by developmental changes in numerical magnitude representations and/or the access to these representations or whether the null effect in children mainly resulted from the dual-task situation and the associated compromise in WM resources especially in the latter individuals.

AUTHOR CONTRIBUTIONS

Conceived and designed the experiments: A-MS, CtS, ClS, and MG. Analyzed the data: CG. Wrote the paper: CG, CtS, and ClS.

FUNDING

The current research was supported by the National Research Fund Luxembourg (FNR; www.fnr.lu) under Grant AFR PhD-2013-1/5558196.

REFERENCES

- Abrahamse, E., van Dijck, J.-P., and Fias, W. (2016). How does working memory enable number-induced spatial biases? *Front. Psychol.* 7:777. doi: 10.3389/fpsyg.2016.00977
- Anobile, G., Cicchini, G. M., and Burr, D. C. (2012). Linear mapping of numbers onto space requires attention. *Cognition* 122, 454–459. doi: 10.1016/j.cognition.2011.11.006
- Baddeley, A. (1998). Random generation and the executive control of working memory. *Q. J. Exp. Psychol. A* 51, 819–852. doi: 10.1080/713755788
- Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645. doi: 10.1146/annurev.psych.59.103006.093639
- Barth, H. C., and Paladino, A. M. (2011). The development of numerical estimation: evidence against a representational shift. *Dev. Sci.* 14, 125–135. doi: 10.1111/j.1467-7687.2010.00962.x
- Berch, D. B., Foley, E. J., Hill, R. J., and Ryan, P. M. (1999). Extracting parity and magnitude from Arabic numerals: developmental changes in number processing and mental representation. *J. Exp. Child Psychol.* 74, 286–308. doi: 10.1006/jecp.1999.2518
- Best, J. R., Miller, P. H., and Jones, L. L. (2009). Executive functions after age 5: changes and correlates. *Dev. Rev.* 29, 180–200. doi: 10.1016/j.dr.2009.05.002
- Booth, J. L., and Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Dev. Psychol.* 42, 189–201. doi: 10.1037/0012-1649.41.6.189
- Brugger, P. (1997). Variables that influence the generation of random sequences: an update. *Percept. Mot. Skills* 84, 627–661. doi: 10.2466/pms.1997.84.2.627
- Cheng, X., Ge, H., Andoni, D., Ding, X., and Fan, Z. (2015). Composite body movements modulate numerical cognition: evidence from the motion-numerical compatibility effect. *Front. Psychol.* 6:1692. doi: 10.3389/fpsyg.2015.01692
- Cohen, D. J., and Blanc-Goldhammer, D. (2011). Numerical bias in bounded and unbounded number line tasks. *Psychon. Bull. Rev.* 18, 331–338. doi: 10.3758/s13423-011-0059-z
- Crollen, V., Dormal, G., Seron, X., Lepore, F., and Collignon, O. (2013). Embodied numbers: the role of vision in the development of number-space interactions. *Cortex* 49, 276–283. doi: 10.1016/j.cortex.2011.11.006
- Crollen, V., and Noël, M.-P. (2015). Spatial and numerical processing in children with high and low visuospatial abilities. *J. Exp. Child Psychol.* 132(Suppl. C), 84–98. doi: 10.1016/j.jecp.2014.12.006
- Crollen, V., Vanderclausen, C., Allaire, F., Pollaris, A., and Noël, M.-P. (2015). Spatial and numerical processing in children with non-verbal learning disabilities. *Res. Dev. Disabil.* 47(Suppl. C), 61–72. doi: 10.1016/j.ridd.2015.08.013
- de Hevia, M. D., Girelli, L., Addabbo, M., and Macchi Cassia, V. (2014a). Human infants' preference for left-to-right oriented increasing numerical sequences. *PLoS One* 9:e96412. doi: 10.1371/journal.pone.0096412
- de Hevia, M. D., Girelli, L., and Vallar, G. (2006). Numbers and space: a cognitive illusion? *Exp. Brain Res.* 168, 254–264. doi: 10.1007/s00221-005-0084-0
- de Hevia, M. D., Izard, V., Coubart, A., Spelke, E. S., and Streri, A. (2014b). Representations of space, time, and number in neonates. *Proc. Natl. Acad. Sci. U.S.A.* 111, 4809–4813. doi: 10.1073/pnas.1323628111
- de Hevia, M. D., and Spelke, E. S. (2009). Spontaneous mapping of number and space in adults and young children. *Cognition* 110, 198–207. doi: 10.1016/j.cognition.2008.11.003
- de Hevia, M. D., and Spelke, E. S. (2010). Number-space mapping in human infants. *Psychol. Sci.* 21, 653–660. doi: 10.1177/0956797610366091
- De Luca, C. R., Wood, S. J., Anderson, V., Buchanan, J. A., Proffitt, T. M., Mahony, K., et al. (2003). Normative data from the Cantab. I: Development

ACKNOWLEDGMENTS

Some minor content was adapted from Georges' dissertation thesis, defended on 17th February 2017 in Luxembourg (Georges, 2017).

- of executive function over the lifespan. *J. Clin. Exp. Neuropsychol.* 25, 242–254. doi: 10.1076/jcen.25.2.242.13639
- Dehaene, S. (1997). *The Number Sense*. New York, NY: Oxford University Press.
- Dehaene, S. (2005). “Evolution of human cortical circuits for reading and arithmetic: the “neuronal recycling” hypothesis,” in *From Monkey Brain to Human Brain*, eds S. Dehaene, J. R. Duhamel, M. D. Hauser, and G. Rizzolatti (Cambridge, MA: MIT Press), 133–158.
- Dehaene, S., Bossini, S., and Giraux, P. (1993). The mental representation of parity and number magnitude. *J. Exp. Psychol. Gen.* 122, 371–396. doi: 10.1037/0096-3445.122.3.371
- Dehaene, S., and Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron* 56, 384–398. doi: 10.1016/j.neuron.2007.10.004
- Ebersbach, M., Luwel, K., Frick, A., Onghena, P., and Verschaffel, L. (2008). The relationship between the shape of the mental number line and familiarity with numbers in 5- to 9-year old children: Evidence for a segmented linear model. *J. Exp. Child Psychol.* 99, 1–17. doi: 10.1016/j.jecp.2007.08.006
- Eerland, A., Guadalupe, T. M., and Zwaan, R. A. (2011). Leaning to the left makes the eiffel tower seem smaller: posture-modulated estimation. *Psychol. Sci.* 22, 1511–1514. doi: 10.1177/0956797611420731
- Evans, F. J. (1978). Monitoring attention deployment by random number generation: an index to measure subjective randomness. *Psychon. Bull. Rev.* 12, 35–38. doi: 10.3758/BF03329617
- Fias, W., and van Dijck, J.-P. (2016). The temporary nature of number-space interactions. *Can. J. Exp. Psychol.* 70, 33–40. doi: 10.1037/cep0000071
- Fischer, M. H. (2008). Finger counting habits modulate spatial-numerical associations. *Cortex* 44, 386–392. doi: 10.1016/j.cortex.2007.08.004
- Fischer, M. H. (2012). A hierarchical view of grounded, embodied, and situated numerical cognition. *Cogn. Process.* 13, 161–164. doi: 10.1007/s10339-012-0477-5
- Fischer, M. H., and Brugger, P. (2011). When digits help digits: spatial-numerical associations point to finger counting as prime example of embodied cognition. *Front. Psychol.* 2:260. doi: 10.3389/fpsyg.2011.00260
- Fischer, M. H., and Shaki, S. (2014). Spatial associations in numerical cognition—from single digits to arithmetic. *Q. J. Exp. Psychol.* 67, 1461–1483. doi: 10.1080/17470218.2014.927515
- Fry, A. F., and Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biol. Psychol.* 54, 1–34. doi: 10.1016/S0301-0511(00)00051-X
- Gathercole, S. E., Pickering, S. J., Ambridge, B., and Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Dev. Psychol.* 40, 177–190. doi: 10.1037/0012-1649.40.2.177
- Georges, C. (2017). *Number-Space Associations as Indexed by the SNARC Effect – Their Relations to Mathematical Abilities and Anxiety & Their Underlying Cognitive Mechanisms*. Ph.D. thesis, University of Luxembourg, Luxembourg.
- Georges, C., Hoffmann, D., and Schiltz, C. (2016). How math anxiety relates to number-space associations. *Front. Psychol.* 7:1401. doi: 10.3389/fpsyg.2016.01401
- Ginsburg, V., van Dijck, J.-P., Previtali, P., Fias, W., and Gevers, W. (2014). The impact of verbal working memory on number-space associations. *J. Exp. Psychol. Learn. Mem. Cogn.* 40, 976–986. doi: 10.1037/a0036378
- Göbel, S. M., Maier, C. A., and Shaki, S. (2015). Which numbers do you have in mind? Number generation is influenced by reading direction. *Cogn. Process.* 16, 241–244. doi: 10.1007/s10339-015-0715-8
- Göbel, S. M., McCrink, K., Fischer, M. H., and Shaki, S. (2017). Observation of directional storybook reading influences young children's counting direction. *J. Exp. Child Psychol.* 166(Suppl. C), 49–66. doi: 10.1016/j.jecp.2017.08.001
- Hamdan, A. C., de Souza, J. A., and Bueno, O. F. A. (2004). Performance of university students on random number generation at different rates to

- p>
evaluate executive functions.
- Arq. Neuropsiquiatr.*
- 62, 58–60. doi: 10.1590/S0004-282X2004000100010
- Hartmann, M., Grabherr, L., and Mast, F. W. (2012). Moving along the mental number line: interactions between whole-body motion and numerical cognition. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 1416–1427. doi: 10.1037/a0026706
- Herrera, A., Macizo, P., and Semenza, C. (2008). The role of working memory in the association between number magnitude and space. *Acta Psychol.* 128, 225–237. doi: 10.16/j.actpsy.2008.01.002
- Hoffmann, D., Hornung, C., Martin, R., and Schiltz, C. (2013). Developing number–space associations: SNARC effects using a color discrimination task in 5-year-olds. *J. Exp. Child Psychol.* 116, 775–791. doi: 10.1016/j.jecp.2013.07.013
- Hoffmann, D., Mussolin, C., Martin, R., and Schiltz, C. (2014). The impact of mathematical proficiency on the number–space association. *PLoS One* 9:e85048. doi: 10.1371/journal.pone.0085048
- Hubbard, E. M., Piazza, M., Pinel, P., and Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nat. Rev. Neurosci.* 6, 435–448. doi: 10.1038/nrn1684
- Hubbard, E. M., Piazza, M., Pinel, P., and Dehaene, S. (2009). “Numerical and spatial intuitions: a role for posterior parietal cortex?” in *Cognitive Biology: Evolutionary, and Developmental Perspectives on Mind, Brain, and Behavior* eds L. Tommasi, L. Nadel, and M. A. Peterson (Cambridge, MA: MIT press), 221–246.
- Jahanshahi, M., Profice, P., Brown, R. G., Ridding, M. C., Dirnberger, G., and Rothwell, J. C. (1998). The effects of transcranial magnetic stimulation over the dorsolateral prefrontal cortex on suppression of habitual counting during random number generation. *Brain* 121, 1533–1544. doi: 10.1093/brain/121.8.1533
- Lachmair, M., Dudschig, C., De la Vega, I., and Kaup, B. (2014). Relating numeric cognition and language processing: Do numbers and words share a common representational platform? *Acta Psychol.* 148, 107–114. doi: 10.1016/j.actpsy.2013.12.004
- Lakoff, G., and Nunez, R. (2000). *Where Mathematics Comes From: How the Embodied Mind Brings Mathematics into Being*. New York, NY: Basic Books.
- Laski, E. V., and Siegler, R. S. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child Dev.* 78, 1723–1743. doi: 10.1111/j.1467-8624.2007.01087.x
- Loetscher, T., Bockisch, C. J., Nicholls, M. E., and Brugger, P. (2010). Eye position predicts what number you have in mind. *Curr. Biol.* 20, R264–R265. doi: 10.1016/j.cub.2010.01.015
- Loetscher, T., Schwarz, U., Schubiger, M., and Brugger, P. (2008). Head turns bias the brain’s internal random generator. *Curr. Biol.* 18, R60–R62. doi: 10.1016/j.cub.2007.11.015
- Lourenco, S. F., and Longo, M. R. (2010). General magnitude representation in human infants. *Psychol. Sci.* 21, 873–881. doi: 10.1177/0956797610370158
- Luciana, M., and Nelson, C. A. (1998). The functional emergence of prefrontally-guided working memory systems in four- to eight-year-old children. *Neuropsychologia* 36, 273–293. doi: 10.1016/S0028-3932(97)00109-7
- Moeller, K., Pixner, S., Kaufmann, L., and Nuerk, H.-C. (2009). Children’ early mental number line: logarithmic or decomposed linear. *J. Exp. Child Psychol.* 103, 503–515. doi: 10.1016/j.jecp.2009.02.006
- Nava, E., Rinaldi, L., Bulf, H., and Macchi Cassia, V. (2017). Visual and proprioceptive feedback differently modulate the spatial representation of number and time in children. *J. Exp. Child Psychol.* 161(Suppl. C), 161–177. doi: 10.1016/j.jecp.2017.04.012
- Ninaus, M., Moeller, K., Kaufmann, L., Fischer, M. H., Nuerk, H.-C., and Wood, G. (2017). Cognitive mechanisms underlying directional and non-directional spatial-numerical associations across the lifespan. *Front. Psychol.* 8:1421. doi: 10.3389/fpsyg.2017.01421
- Opfer, J. E., and Furlong, E. E. (2011). How numbers bias preschoolers’ spatial search. *J. Cross Cultural Psychol.* 42, 682–695. doi: 10.1177/0022022111406098
- Opfer, J. E., and Siegler, R. S. (2007). Representational change and children’s numerical estimation. *Cogn. Psychol.* 55, 169–195. doi: 10.1016/j.cogpsych.2006.09.002
- Opfer, J. E., Thompson, C. A., and Furlong, E. E. (2010). Early development of spatial-numeric associations: evidence from spatial and quantitative performance of preschoolers. *Dev. Sci.* 13, 761–771. doi: 10.1111/j.1467-7687.2009.00934.x
- Patro, K., and Haman, M. (2012). The spatial–numerical congruity effect in preschoolers. *J. Exp. Child Psychol.* 111, 534–542. doi: 10.1016/j.jecp.2011.09.006
- Peters, M., Giesbrecht, T., Jelacic, M., and Merckelbach, H. (2007). The random number generation task: psychometric properties and normative data of an executive function task in a mixed sample. *J. Int. Neuropsychol. Soc.* 13, 626–634. doi: 10.1017/S1355617707070786
- Shaki, S., and Fischer, M. H. (2014). Random walks on the mental number line. *Exp. Brain Res.* 232, 43–49. doi: 10.1007/s00221-013-3718-7
- Shaki, S., Fischer, M. H., and Göbel, S. M. (2012). Direction counts: a comparative study of spatially directional counting biases in cultures with different reading directions. *J. Exp. Child Psychol.* 112, 275–281. doi: 10.1016/j.jecp.2011.12.005
- Shaki, S., Fischer, M. H., and Petrusic, W. M. (2009). Reading habits for both words and numbers contribute to the SNARC effect. *Psychon. Bull. Rev.* 16, 328–331. doi: 10.3758/PBR.16.2.328
- Siegler, R., and Booth, J. (2004). Development of numerical estimation in young children. *Child Dev.* 75, 428–444. doi: 10.1111/j.1467-8624.2004.00684.x
- Siegler, R. S., and Opfer, J. E. (2003). The development of numerical estimation: evidence for multiple representations of numerical quantity. *Psychol. Sci.* 14, 237–243. doi: 10.1111/1467-9280.02438
- Simms, V., Clayton, S., Cragg, L., Gilmore, C., and Johnson, S. (2016). Explaining the relationship between number line estimation and mathematical achievement: the role of visuomotor integration and visuospatial skills. *J. Exp. Child Psychol.* 145, 22–33. doi: 10.1016/j.jecp.2015.12.004
- Slusser, E. B., Santiago, R. T., and Barth, H. C. (2013). Developmental change in numerical estimation. *J. Exp. Psychol. Gen.* 142, 193–208. doi: 10.1037/a0028560
- Terhune, D. B., and Brugger, P. (2011). Doing better by getting worse: posthypnotic amnesia improves random number generation. *PLoS One* 6:e29206. doi: 10.1371/journal.pone.0029206
- Thompson, J. M., Nuerk, H.-C., Moeller, K., and Cohen Kadosh, R. (2013). The link between mental rotation ability and basic numerical representations. *Acta Psychol.* 144, 324–331. doi: 10.1016/j.actpsy.2013.05.009
- Towse, J. N., and Cheshire, A. (2007). Random number generation and working memory. *Eur. J. Cogn. Psychol.* 19, 374–394. doi: 10.1080/09541440600764570
- Towse, J. N., Loetscher, T., and Brugger, P. (2014). Not all numbers are equal: preferences and biases among children and adults when generating random sequences. *Front. Psychol.* 5:19. doi: 10.3389/fpsyg.2014.00019
- Towse, J. N., and Neil, D. (1998). Analyzing human random generation behavior: a review of methods used and a computer program for describing performance. *Behav. Res. Methods Instrum. Comput.* 30, 583–591. doi: 10.3758/BF03209475
- Ullman, H., Almeida, R., and Klingberg, T. (2014). Structural maturation and brain activity predict future working memory capacity during childhood development. *J. Neurosci.* 34, 1592–1598. doi: 10.1523/JNEUROSCI.0842-13.2014
- van Dijck, J.-P., Abrahamse, E. L., Acar, F., Ketels, B., and Fias, W. (2014). A working memory account of the interaction between numbers and spatial attention. *Q. J. Exp. Psychol.* 67, 1500–1513. doi: 10.1080/17470218.2014.903984
- van Dijck, J.-P., and Fias, W. (2011). A working memory account for spatial–numerical associations. *Cognition* 119, 114–119. doi: 10.1016/j.cognition.2010.12.013
- van Dijck, J.-P., Gevers, W., and Fias, W. (2009). Numbers are associated with different types of spatial information depending on the task. *Cognition* 113, 248–253. doi: 10.1016/j.cognition.2009.08.005
- van Galen, M. S., and Reitsma, P. (2008). Developing access to number magnitude: a study of the SNARC effect in 7- to 9-year-olds. *J. Exp. Child Psychol.* 101, 99–113. doi: 10.1016/j.jecp.2008.05.001
- White, S. L., Szűcs, D., and Soltész, F. (2012). Symbolic number: the integration of magnitude and spatial representations in children aged 6 to 8 years. *Front. Psychol.* 2:392. doi: 10.3389/fpsyg.2011.00392
- Winter, B., and Matlock, T. (2013). “More is up and right: random number generation along two axes,” in *Proceedings of the 35th Annual*

- Conference of the Cognitive Science Society*, eds M. Knauff, M. Pauen, N. Sebanz, and I. Wachsmuth (Austin, TX: Cognitive Science Society), 3789–3974.
- Yang, T., Chen, C., Zhou, X., Jihong, X., Dong, Q., and Chen, C. (2014). Development of spatial representation of numbers: a study of the SNARC effect in Chinese children. *J. Exp. Child Psychol.* 117, 1–11. doi: 10.1016/j.jecp.2013.08.11
- Zebian, S. (2005). Linkages between number concepts, spatial thinking, and directionality of writing: the SNARC effect and the reverse SNARC effect in English and Arabic monoliterates, Biliterates, and illiterate Arabic speakers. *J. Cogn. Cult.* 5, 165–190. doi: 10.1163/1568537054068660
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Sosson, Georges, Guillaume, Schuller and Schiltz. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



From Innate Spatial Biases to Enculturated Spatial Cognition: The Case of Spatial Associations in Number and Other Sequences

Koleen McCrink^{1*} and Maria Dolores de Hevia^{2,3}

¹ Department of Psychology, Barnard College, Columbia University, New York, NY, United States, ² Université Paris Descartes, Paris, France, ³ Laboratoire Psychologie de la Perception, CNRS UMR 8242, Paris, France

Keywords: space, number, neonates, laterality, toddlerhood

OPEN ACCESS

Edited by:

Hans-Christoph Nuerk,
Universität Tübingen, Germany

Reviewed by:

Wim Fias,
Ghent University, Belgium
Elizabeth M. Brannon,
Duke University, United States
Vanessa R. Simmering,
University of Wisconsin-Madison,
United States

*Correspondence:

Koleen McCrink
kmcCrink@barnard.edu

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 20 November 2017

Accepted: 13 March 2018

Published: 29 March 2018

Citation:

McCrink K and de Hevia MD (2018)
From Innate Spatial Biases to
Enculturated Spatial Cognition: The
Case of Spatial Associations in
Number and Other Sequences.
Front. Psychol. 9:415.
doi: 10.3389/fpsyg.2018.00415

Humans, as well as other animals, use space to organize the world. This use of space as an organizational scaffold is especially prevalent when we conceptualize mathematics, a domain that shares behavioral and neural overlap with the domain of space (Pinel et al., 2004; Kaufmann et al., 2005; Dehaene and Brannon, 2010). One of the most prominent descriptions of this relation is that of a mental number line, in which small values are associated with the left side of space, and large values with the right (Moyer and Landauer, 1967; Dehaene et al., 1993). The development of the mature form of this mental number line is multiply determined, with evidence pointing to evolutionary pressures as well as cultural and linguistic influences. This cognitive bias to associate numerical information with space, and do so with left-right or right-left asymmetry, is adaptive; it helps to bolster memory and learning throughout our lives (Opfer and Furlong, 2011; McCrink and Galamba, 2015; McCrink and Shaki, 2016; Bulf et al., 2017). Moreover, with development this bias to map number onto an oriented continuum extends to any well-ordered information, even when recently learned (Gevers et al., 2003, 2004; Previtali et al., 2010). Critically, despite the apparent promise of using space as a scaffold for learning and memory, there are several gaps in the literature surrounding an essential period of the development of spatial-numerical associations: toddlerhood and early childhood. Here, we summarize current work on the innate and culture-specific factors modulating the mental number line in infancy and childhood, and note further research that could help to shed light on a complete developmental picture of this phenomenon.

THE MENTAL NUMBER LINE: FROM INNATE TO ENCULTURATED

Recent work in developmental psychology has found that spatial-numerical associations are present as early as the first days of life. de Hevia and colleagues have documented a propensity for infants in the first year of life to map magnitudes onto a left-to-right spatial continuum. Seven-month old infants present a preference for increasing numerical sequences, only if the arrays are presented from smallest on the left to largest on the right (de Hevia et al., 2014). Eight-month-olds are quicker to attend to a left-side probe after central presentation of a small number and a right-side probe after central presentation of a large number, but this advantage does not extend to a small vs. large object (Bulf et al., 2016). Interestingly, despite numerical magnitude and spatial quantity sharing many commonalities in infancy [e.g., an advantage for increasing order (Macchi Cassia et al., 2012; de Hevia et al., 2014, 2017a), transfer of ordinal direction and rule-based learning between the two domains (de Hevia and Spelke, 2010; Lourenco and Longo, 2010)], the findings of lateralized asymmetry for attention in infancy seem to be specific to numerical magnitude (e.g., sets of objects) and not spatial quantity (e.g., the size of a single object; Bulf et al., 2016; de Hevia et al., 2017b). This lateralized processing can be found even when the dimension evokes number only peripherally, such as when processing a statistical ordering rule for the placement of three objects (Bulf et al., 2017). The biases observed in infancy are untrained and spontaneous, reflecting predispositions for lateralized processing of magnitude. However, it is possible that by several months of age, infants

have had some non-specific spatial experience that could lead to enculturation of a spatial organization system. de Hevia et al. (2017a) have recently found that even neonates exhibit lateralized processing of magnitude; they look longer to a left-side stimulus in the presence of a relatively small magnitude, and longer to a right-side stimulus in the presence of a relatively large magnitude. This finding—which is not mutually exclusive with a later, enculturated mental number line—supports the existence of a mental number line in humans with no prior spatial experience.

McCrink et al. (2017b) posited that these lateralized spatial-numerical associations wax and wane throughout infancy and early childhood as children become less beholden to innate biases, and more imitative and aware of the cultural conventions surrounding spatial structuring. In this study, 2- and 3-year-olds were given a version of a navigational spatial transposition task frequently used with non-human animals (Rugani et al., 2010; Drucker and Brannon, 2014). In the experimental conditions relevant to this review, toddlers were trained to retrieve an object that was repeatedly hidden in one particular location (out of 5) along a vertical array, with the experimenter verbally labeling the locations with numerals (“box one”) or a non-ordinal label (“this box”). Afterwards, the array was surreptitiously transposed 90 degrees. Unlike non-human animals, who exhibit a general bias to search from left-to-right after being trained in this spatially ordered sequence of locations, the children who received generic labels were equally likely to navigate with a LR or RL bias. However, children who received numerical labels selected the location that corresponded to a left-to-right spatial mapping.

Moreover, in a counting task only ~60% of toddlers counted in an organized direction, and those were the children who reliably performed a left-to-right mapping. In light of these findings, the authors suggest that toddlerhood is a period of flexibility with respect to the directional nature of spatial associations, with innate left-to-right scanning biases falling away as children begin to gather socially transmitted information of the spatial structuring in their environment. Early biases to map initial information to the left side of space, and final to the right, will arise only if the privileged domain of number is invoked (See **Figure 1** for the proposed developmental trajectory of several types of spatial associations).

This privileged mapping of numerals to space is likely due to the combination of the children’s knowledge of the mapping between numerals and magnitude (an inherently ordinal dimension), and the reinforcement of left-to-right spatial structuring by their caregivers when counting. During the preschool years, children start to reliably map small numbers (“1, 2, 3”) to their innate, non-symbolic, and intrinsically ordered representations of number (Sarnecka and Carey, 2008). By preschool, children show spatial-numerical compatibility effects similar to older children and adults for non-symbolic magnitudes (de Hevia and Spelke, 2009; Patro and Haman, 2012), and are more likely to use symbolic numerical labels to solve a spatial reasoning task if they are presented in a culturally consistent direction (Opfer et al., 2010). In this paradigm (adapted from Loewenstein and Gentner, 2005), preschoolers are shown two sets of boxes (a sample and matching set), sectioned into verbally

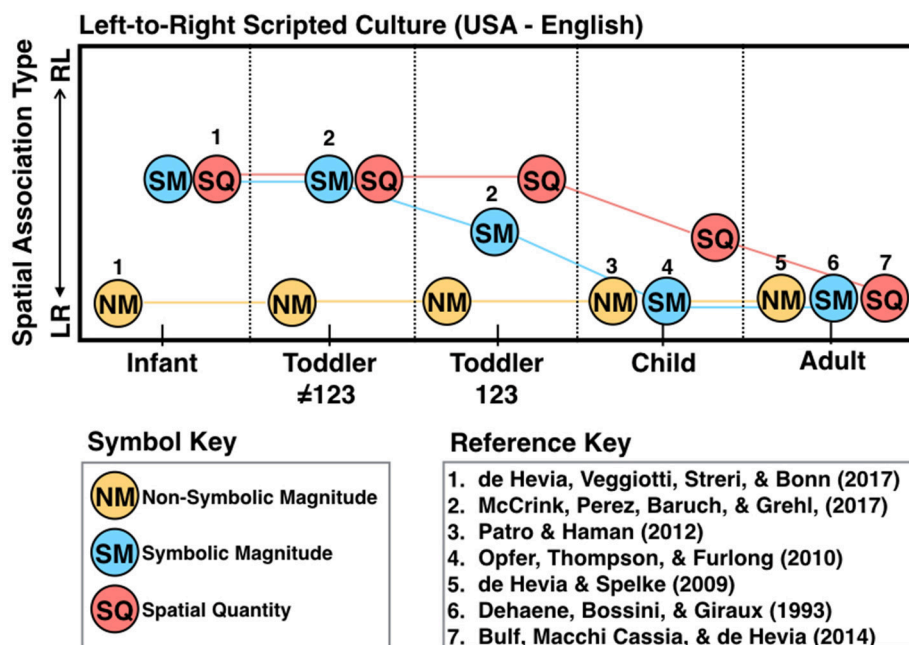


FIGURE 1 | A summary of likely trajectories of the spatial association types [ranging from left-small/less and right-large /more (LR) to right-small/less and left-large /more (RL)] in early childhood, for a child whose language is consistently scripted left to right. Knowledge of the count list is indicated here by “123.” Numbers below the symbols indicate the studies which have been done to establish this trajectory, with author details noted on the reference key. A lack of numbers indicates an area for future work. In infancy, children spontaneously associate small and large magnitudes with the left and right sides of space (respectively). In the toddler years, children develop symbolic knowledge of the order of numerals, and are enculturated to the different spatial structures within their script for these symbols, which eventually prompts culture-specific spatial associations for many types of ordered sequences.

labeled locations (e.g., “room 2”). A target is shown in the sample set, and children search for this target in the matching set (located in the same labeled location). Preschoolers in the U.S. are faster and more accurate when locations are numbered from left-to-right versus right-to-left, if they are highly organized counters (Opfer et al., 2010). Additionally, Shaki et al. (2012) found that preschoolers in cultures with right-to-left scripted language (such as Arabic) exhibit spatial-numerical biases that are reversed, with young children counting from right-to-left instead of from left-to-right as they do in English-speaking countries.

How may this conventionality emerge? Given the timing of this shift, the obvious candidate is the child's home environment. Starting in early toddlerhood, caregivers are modeling the spatial conventions of their culture, presenting spatial associations with a high degree of culture-specific structure. Parents may primarily model a single effective strategy when they organize space for their child—a strategy that is colored by the language they read and write on a daily basis. Recent work on caregiving influences on spatial biases suggests there are three primary ways that parents can influence their child's spatial structuring habits: their gesture, their organization of spatial layout, and the nature of their reading material (Patro et al., 2016a; Göbel et al., 2017; McCrink et al., 2017a). McCrink et al. (2017a) found that in two different tasks—watching a slideshow of alphabetical, numerical, or random stimuli, and crafting a visual story for their child—English-speaking parents were more likely to gesture to the screen and lay out pictures in a left-to-right manner to a greater degree than Hebrew-speaking parents. Göbel et al. (2017) found that after observing reading from storybooks (a left-to-right or right-to-left storybook) children change their counting direction in line with the direction of reading. Observing an adult point in a specific direction (e.g., right to left) did not influence counting direction. In contrast, Patro et al. (2016b) found that if the children were trained by an adult to point in a specific direction themselves, their subsequent spatial-numerical mappings took on the asymmetric form of that pointing movement (left-less/right-more after left-to-right pointing, and right-less/left-more after right-to-left pointing). Finally, book illustrations exhibit culture-specific directionality, even in non-numerical domains, with the subject[object] of the sentence on the left[right] for English-language books, and the opposite for Hebrew-language books (Göbel et al., 2017). The accumulation of this cultural experience results in an asymmetric mapping for many types of ordinal information (numerical: Dehaene et al., 1993; Zebian, 2005, spatial quantity: Bulf et al., 2014, alphabetical: McCrink and Shaki, 2016)—a mapping which follows the direction of the culture's script.

FUTURE DIRECTIONS ON THE EARLY DEVELOPMENT OF THE MENTAL NUMBER LINE

Several outstanding questions remain within this subfield. First, is the number-space mapping in infancy actually related to the

ubiquitous spatial associations found in adulthood? It is instead possible that these are two separate phenomena, which reflect different underlying mechanisms [e.g., hemispheric lateralization influences in infancy, but a distinct symbolic, analogical reasoning system starting in the second year of life Halford et al., 2010, 2013]. One way to address this possibility is to investigate both the structure and function of brain areas which respond to numerical and spatial magnitudes (e.g., Borghesani et al., 2016), and observe if there is continuity across development with respect to which regions are activated in similar tasks. Second, what is the underlying spatial relation between different types of quantity representations at birth? Studies which investigate the numerical specificity of spatial associations in neonates should be conducted in order to detail how the domain of number is structured and reasoned about. Third, when does the enculturation shift for spatial associations happen—and does the presence or absence of numerical input alter this timeline? To answer this question, research is needed in which the same spatial association task is implemented in infants, toddlers, and children in cultures which observe left-to-right and right-to-left scripting behaviors. One good candidate would be the spatial transposition task, which requires no verbal knowledge, and can be altered for the presence or absence of non-symbolic number arrays on each location. Fourth, how exactly is this enculturation of spatial associations implemented? Work on spatial enculturation behaviors like gesturing along a path (Patro et al., 2016b) and reading (Göbel et al., 2017) has started to document possible avenues, but a closer study of the home environment and the relation between parent behaviors and child spatial associations is needed. For example, if reading observation is a primary avenue to enculturation for this phenomenon, one would predict that highly literate homes would have children who exhibit a quicker and more robust transition to the spatial associations of their culture. Additionally, a causal story for parent interaction as the driver of enculturated spatial associations would predict that parents' degree of spatial structuring would be the modulating factor in their child's degree of spatial associations. Finally, the relation between different types of enculturation behaviors and different types of numerical representations is still unclear. Developmental studies which systematically tease apart the influence of these behaviors (a parent modeling spatial organization vs. a child mimicking these modeled behaviors, parental modeling of spatial organization in a numerical or non-numerical fashion) and representations (explicit counting, non-symbolic mapping of magnitudes) could help clarify the nature of the mental number line in early childhood.

AUTHOR CONTRIBUTIONS

KM and MDdH contributed equally to the generation of this opinion. KM drafted the manuscript. MDdH provided comments.

FUNDING

This manuscript was supported by funds from the Eunice Kennedy Shriver National Institute of Child Health and

Human Development (R15 HD077518-01A1) to KM. This research was supported by an ANR (Agence National de la Recherche Scientifique ANR-15-CE28-0003-01 NUMSPA) to MDdH.

REFERENCES

- Borghesani, V., de Hevia, M. D., Viarouge, A., Chagas, P. P., Eger, E., and Piazza, M. (2016). "Comparing magnitudes across dimensions: a univariate and multivariate approach," in *Pattern Recognition in Neuroimaging (PRNI), 2016 International Workshop* (Trento: IEEE).
- Bulf, H., de Hevia, M. D., Gariboldi, V., and Macchi Cassia, V. (2017). Infants learn better from left to right: a directional bias in infants' sequence learning. *Sci. Rep.* 7:2437. doi: 10.1038/s41598-017-02466-w
- Bulf, H., de Hevia, M. D., and Macchi Cassia, V. (2016). Small on the left, large on the right: numbers orient visual attention onto space in preverbal infants. *Dev. Sci.* 19, 394–401. doi: 10.1111/desc.12315
- Bulf, H., Macchi Cassia, V., and de Hevia, M. D. (2014). Are numbers, brightness, and size equally efficient in orienting visual attention? Evidence from an eye tracking study. *PLoS ONE* 9:e99499. doi: 10.1371/journal.pone.0099499
- Dehaene, S., Bossini, S., and Giraux, P. (1993). The mental representation of parity and number magnitude. *J. Exp. Psychol. Gen.* 122, 371–396. doi: 10.1037/0096-3445.122.3.371
- Dehaene, S., and Brannon, E. (2010). *Space, Time and Number in the Brain: Searching for the Foundations of Mathematical Thought*. Cambridge: Academic Press.
- de Hevia, M. D., Addabbo, M., Nava, E., Croci, E., Girelli, L., and Macchi Cassia, V. (2017a). Infants' detection of increasing numerical order comes before detection of decreasing number. *Cognition* 158, 177–188. doi: 10.1016/j.cognition.2016.10.022
- de Hevia, M. D., Girelli, L., Addabbo, M., and Macchi Cassia, V. (2014). Human infants' preference for left-to-right oriented increasing numerical sequences. *PLoS ONE* 9:e96412. doi: 10.1371/journal.pone.0096412
- de Hevia, M. D., and Spelke, E. S. (2009). Spontaneous mapping of number and space in adults and young children. *Cognition* 110, 198–207. doi: 10.1016/j.cognition.2008.11.003
- de Hevia, M. D., and Spelke, E. S. (2010). Number-space mapping in human infants. *Psychol. Sci.* 21, 653–660. doi: 10.1177/0956797610366091
- de Hevia, M. D., Veggioni, L., Streri, A., and Bonn, C. (2017b). At birth human infants associate 'few' with left and 'many' with right. *Curr. Biol.* 27, 3879–3884. doi: 10.1016/j.cub.2017.11.024
- Drucker, C., and Brannon, E. (2014). Rhesus monkeys (*Macaca mulatta*) map number onto space. *Cognition* 132, 57–67. doi: 10.1016/j.cognition.2014.03.011
- Gevers, W., Reynvoet, B., and Fias, W. (2003). The mental representation of ordinal sequences is spatially organized. *Cognition* 87, B87–B95. doi: 10.1016/S0010-0277(02)00234-2
- Gevers, W., Reynvoet, B., and Fias, W. (2004). The mental representation of ordinal sequences is spatially organized: evidence from days of the week. *Cortex* 40, 171–172. doi: 10.1016/S0010-9452(08)70938-9
- Göbel, S., McCrink, K., Fischer, M., and Shaki, S. (2017). Observation of directional storybook reading influences young children's counting direction. *J. Exp. Child Psychol.* 166, 49–66. doi: 10.1016/j.jecp.2017.08.001
- Halford, G., Andrews, G., Phillips, S., and Wilson, W. (2013). The role of working memory in the subsymbolic-symbolic transition. *Curr. Dir. Psychol. Sci.* 22, 210–216. doi: 10.1177/0963721412470132
- Halford, G., Wilson, W., and Phillips, S. (2010). Relational knowledge: the foundation of higher cognition. *Trends Cogn. Sci.* 14, 497–505. doi: 10.1016/j.tics.2010.08.005
- Kaufmann, L., Koppelstaetter, F., Delazer, M., Siedentopf, C., Rhomberg, P., Golaszewski, S., et al. (2005). Neural correlates of distance and congruity effects in a numerical Stroop task: an event-related fMRI study. *Neuroimage* 25, 888–898. doi: 10.1016/j.neuroimage.2004.12.041
- Loewenstein, J., and Gentner, D. (2005). Relational language and the development of relational mapping. *Cogn. Psychol.* 50, 315–353. doi: 10.1016/j.cogpsych.2004.09.004
- Lourenco, S. F., and Longo, M. R. (2010). General magnitude representation in human infants. *Psychol. Sci.* 21, 873–881. doi: 10.1177/0956797610370158
- Macchi Cassia, V., Picozzi, M., Girelli, L., and de Hevia, M. D. (2012). Increasing magnitude counts more: asymmetrical processing of ordinality in 4-month-old infants. *Cognition* 124, 183–193. doi: 10.1016/j.cognition.2012.05.004
- McCrink, K., Caldera, C., and Shaki, S. (2017a). The early construction of spatial attention: culture, space, and gesture in parent-child interactions. *Child Dev.* doi: 10.1111/cdev.12781. [Epub ahead of print].
- McCrink, K., and Galamba, J. (2015). The impact of symbolic and non-symbolic quantity on spatial learning. *PLoS ONE* 10:e0119395. doi: 10.1371/journal.pone.0119395
- McCrink, K., Perez, J., Baruch, E., and Grehl, M. (2017b). Numerical labels prompt left-to-right mapping in toddlers. *Dev. Psychol.* 53, 1256–1264. doi: 10.1037/dev0000342
- McCrink, K., and Shaki, S. (2016). Culturally inconsistent spatial structure reduces learning. *Acta Psychol.* 169, 20–26. doi: 10.1016/j.actpsy.2016.05.007
- Moyer, R. S., and Landauer, T. K. (1967). Time required for judgments of numerical inequality. *Nature* 215, 1519–1520. doi: 10.1038/2151519a0
- Opfer, J. E., and Furlong, E. E. (2011). How numbers bias preschoolers' spatial search. *J. Cross. Cult. Psychol.* 42, 682–695. doi: 10.1177/0022022111406098
- Opfer, J. E., Thompson, C. A., and Furlong, E. E. (2010). Early development of spatial-numeric associations: evidence from spatial and quantitative performance of preschoolers. *Dev. Sci.* 13, 761–771. doi: 10.1111/j.1467-7687.2009.00934.x
- Patro, K., Fischer, U., Nuerk, H., and Cress, U. (2016a). How to rapidly construct a spatial-numerical representation in preliterate children (at least temporarily). *Dev. Sci.* 19, 126–144. doi: 10.1111/desc.12296
- Patro, K., and Haman, M. (2012). The spatial-numerical congruity effect in preschoolers. *J. Exp. Child Psychol.* 111, 534–542. doi: 10.1016/j.jecp.2011.09.006
- Patro, K., Nuerk, H. C., and Cress, U. (2016b). Mental number line in the preliterate brain: the role of early directional experiences. *Child Dev. Perspect.* 10, 172–177. doi: 10.1111/cdep.12179
- Pinel, P., Piazza, M., Le Bihan, D., and Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron* 41, 983–993. doi: 10.1016/S0896-6273(04)00107-2
- Previtali, P., de Hevia, M. D., and Girelli, L. (2010). Placing order in space: the SNARC effect in serial learning. *Exp. Brain Res.* 201, 599–605. doi: 10.1007/s00221-009-2063-3
- Rugani, R., Kelly, D. M., Szelest, I., Regolin, L., and Vallortigara, G. (2010). Is it only humans that count from left to right? *Biol. Lett.* 6, 290–292. doi: 10.1098/rsbl.2009.0960
- Sarnecka, B. W., and Carey, S. (2008). How counting represents number: what children must learn and when they learn it. *Cognition* 108, 662–674. doi: 10.1016/j.cognition.2008.05.007
- Shaki, S., Fischer, M. H., and Göbel, S. M. (2012). Direction counts: a comparative study of spatially directional counting biases in cultures with different reading directions. *J. Exp. Child Psychol.* 112, 275–281. doi: 10.1016/j.jecp.2011.12.005
- Zebian, S. (2005). Linkages between number concepts, spatial thinking, and directionality of writing: the SNARC effect and the REVERSE SNARC effect in English and Arabic monoliterates, biliterates, and illiterate Arabic speakers. *J. Cogn. Cult.* 5, 165–190. doi: 10.1163/1568537054068660

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 McCrink and de Hevia. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Magnitude or Multitude – What Counts?

**Martin Lachmair^{1*}, Susana Ruiz Fernández^{1,2,3}, Korbinian Moeller^{1,3,4},
Hans-Christoph Nuerk^{1,3,4} and Barbara Kaup^{3,4}**

¹ Leibniz-Institut für Wissensmedien, Tübingen, Germany, ² FOM-Hochschule für Oekonomie und Management, Essen, Germany, ³ LEAD Graduate School and Research Network, University of Tübingen, Tübingen, Germany,

⁴ Department of Psychology, University of Tübingen, Tübingen, Germany

Recent studies revealed an association of low or high numbers (e.g., 1 vs. 9) and word semantics referring to entities typically found in upper or lower space (e.g., *roof* vs. *root*) indicating overlapping spatial representations. Another line of research revealed a similar association of grammatical number as a syntactic aspect of language and physical space: singular words were associated with left and plural words with right - resembling spatial-numerical associations of low numbers with left and high numbers with right.

The present study aimed at integrating these lines of research by evaluating both types of spatial relations in one experiment. In a lexical decision task, pairs of a numerical cue and a subsequent plural noun were presented. For word with spatial associations (e.g., *roofs* vs. *roots*) number magnitude was expected to serve as a spatial cue. For spatially neutral words (e.g., *tables*) numbers were expected to cue multitude. Results showed the expected congruency-effect between the numbers and words with spatial associations (i.e., small numbers facilitate responses to down-words and high numbers to up-words). However, no effect was found for numbers and spatially neutral words. This seems to indicate that spatial aspects of word meaning may be related more closely to the magnitude of numbers than grammatical number is to the multitude reflected by numbers – at least in the current experimental setting, where only plural words were presented.

Keywords: numerical cognition, grammatical number, space-number associations, space-word associations, grounded cognition

OPEN ACCESS

Edited by:

Ann Dowker,
University of Oxford, United Kingdom

Reviewed by:

Christine Schiltz,
University of Luxembourg,
Luxembourg
Samuel Shaki,
Ariel University, Israel

*Correspondence:

Martin Lachmair
m.lachmair@iwm-kmrc.de;
m.lachmair@iwm-tuebingen.de

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 15 November 2017

Accepted: 27 March 2018

Published: 12 April 2018

Citation:

Lachmair M, Ruiz Fernández S,
Moeller K, Nuerk H-C and Kaup B
(2018) Magnitude or Multitude – What
Counts? *Front. Psychol.* 9:522.
doi: 10.3389/fpsyg.2018.00522

INTRODUCTION

Human language and human's ability for numerical cognition evolved in the context of the physical conditions on Earth. For example, gravitational force of earth gives us an omnipresent reference of vertical space. Thus, it may come with no surprise that such conditions have shaped human cognitive systems. This, for example, is reflected in human language, which is full of words and phrases that explicitly or implicitly express spatial attributes related to the vertical spatial dimension (cf. Levinson, 2003; Lakoff and Johnson, 2008). In addition, this vertical spatial dimension also plays an important role in numerical cognition (cf. Dehaene, 2011; Fischer and Shaki, 2014 for a review). In cognitive science, important lines of research pursue how information that is captured in such symbolic systems like language and numbers is represented mentally.

In principle, it is possible that such representations are based on abstract, arbitrary and amodal cognitive processes (e.g., Fodor, 1975) that reside within memory systems separate from the brain's modal systems (e.g., perception, action; Tulving, 1972). However, over the last decades there has been accumulating evidence for mental representations based on sensorimotor experience, suggesting an important role of sensorimotor aspects in knowledge representations (cf. Barsalou, 2012).

Spatial Representations as a Common Ground of Words and Numbers

Several studies showed that words may automatically activate spatial information related to the typical location of their referents. For example, a word like “roof” whose referent is typically located and experienced in upper vertical space shifts attention upwards. In contrast, a word like “root” whose typical location is in lower vertical space was observed to shift attention downwards (e.g., Lachmair et al., 2011; Dudschig et al., 2013; Thornton et al., 2013).

For the case of numbers, their dominant and most ubiquitous spatial association is typically referred to by the metaphor of a horizontal mental number line (e.g., Restle, 1970; Dehaene et al., 1993) on which numbers are represented according to their magnitude from left to right (cf. Fischer and Shaki, 2014, for a review). However, many authors considered this unidimensional metaphor as insufficient (e.g., Dehaene, 2011; Cipora et al., 2015; Winter et al., 2015). Interestingly, for such directional spatial-numerical associations, in which a certain direction in space is associated with larger numbers (i.e., right, up, etc.), different dimensions may play a role. For instance, there are also findings suggesting a vertical representation of numbers from lower (small numbers) to upper vertical space (larger numbers; e.g., Schwarz and Keus, 2004). Given that both dimensions are associated with number magnitude, the question arose which spatial dimension (i.e., horizontal or vertical) may be associated more strongly with the representation of number magnitude (cf. Holmes and Lourenco, 2012). In fact, Fischer and Brugger (2011) suggested a hierarchical view of spatial-numerical associations differentiating grounded, embodied and situated aspects in the mental representation of numbers (see also Myachykov et al., 2014). According to this view, the metaphor of a horizontal mental number line is driven by cultural conventions, practices and habits (e.g., left-to-right reading direction) and is therefore considered embodied. In contrast, the vertical representation of numbers was proposed to be grounded in the sense that it is based on and reflects universal physical conditions like gravitational force of earth (Fischer and Brugger, 2011) – and thus be more general than the embodied metaphor of a mental number line. This may be illustrated easily by considering the example of filling a glass with water. As one pours more water into the glass the surface level of the water in the glass rises. This reflects a general grounding experience of more of something (in this case water) being associated spatially and accordingly numerically higher magnitudes regardless of culture or place on earth (cf. Lachmair et al., 2017).

A recent study investigated these strong spatial relationships using numbers and nouns. Lachmair et al. (2014) hypothesized that there may be a common or overlapping representational space for the domains of numbers and words referring to entities typically located in upper or lower vertical space. And indeed, they observed that processing low and high numbers, respectively, affected the processing of subsequent words referring to objects with a typical location in lower or upper vertical space (henceforth referred to as down- and up-words, respectively). In particular, the authors found shorter reaction times in a lexical decision task for a congruent combination of low number primes (e.g., “1,” “2”) and subsequently presented down-words (e.g., “floor”) and high number primes (e.g., “8,” “9”) and subsequently presented up-words (e.g., “sky”). In contrast, reaction times were longer in incongruent combinations of numbers and words (i.e., combinations of low number primes followed by up-words and high number primes followed by down-words). The authors interpreted these results as evidence for an overlap in the meaning representations of numbers and words referring to entities with a typical location in upper vs. lower vertical space (Lachmair et al., 2014). This overlap presumably results from the fact that similar mental states are being activated when interacting with the referents of these two types of symbols in the world (cf. Barsalou, 2012). Thus, according to the above mentioned view proposed by Fischer and Brugger (2011), one may conclude that similar to the grounding of number magnitude on vertical space, attentional shifts subsequent to processing words like “sky” or “floor” also reflect effects of groundedness, because their mental representations integrate experiences according to omnipresent physical conditions.

Embodied or Grounded Spatial Representation of Grammatical Number

Beyond commonalities with respect to spatial attributes of word meaning, words and numbers are also interrelated by the syntactical concept of grammatical number. A recent study by Roettger and Domahs (2015) showed that the flexion of German nouns expressing the multitude of their referent(s) also has a spatial association. The authors found a horizontal spatial association, indicated by faster reaction times for singular words when responded to with the left compared to the right hand, whereas a reversed pattern was observed for words in plural form. Although this pattern was found in relatively late stages of the response process and seemed to vary with the complexity of stimulus decoding, this result indicates that *multitude* derived from the syntactic concept of grammatical number is represented on a horizontal axis with lower quantities (i.e., singular) associated with left and higher quantities (i.e., plural) associated with right.

This raised the question at which level grammatical number and physical space interact. According to the hierarchical structure proposed by Fischer and Brugger (2011, see also Myachykov et al., 2014; Lachmair et al., 2017), the horizontal spatial representation of grammatical number may be considered *embodied* because it relies on an overlearned cultural convention

and not on an omnipresent physical law that may shape human cognition.

However, there is also evidence suggesting a grounded origin of the spatial representation of multitude. In their study, Berent et al. (2005) argued that readers extract syntactic grammatical number of bare nouns automatically and represent it in a way that is comparable to the representation of number they extract from visual stimuli. Thus, one might conclude that identifying quantities is a fundamental and universal ability of the human visuo-spatial perceptual system (Anobile et al., 2016). In turn, this would imply the concept of syntactic grammatical number to be grounded. However, even though this fundamental ability may be invariant across cultures, one may doubt that it is deeply associated with mental representations of grammatical number. If so, one would expect a universal cross-cultural representational system of grammatical number. However, this is obviously not the case when considering languages that differentiate explicitly between singular and plural like English or German on the one and languages that have very little singular/plural marking like Mandarin or Japanese on the other hand (e.g., Downing, 1996; Sarnecka, 2014; see Overmann, 2015 for an overview) or special cases such as some Slavic languages in which grammatical number differs for different number ranges (e.g., Polish). Thus, it is unclear how the mental representation of grammatical number may be embedded in a hierarchical structure as proposed by Fischer and Brugger (2011). It is however, well conceivable that the mental representation of grammatical number is grounded and embodied (and maybe even situated). Which representation is actually accessed may depend on the dimension (horizontal or vertical) in which the relation between grammatical number and numerical magnitude is examined.

Magnitude Versus Multitude

Against this background, the question arises which representation of meaning is affected, multitude or magnitude, when nouns denoting objects in vertical space are presented in plural subsequent to numerical cues. Following the hierarchical view of Fischer and Brugger (2011) one would assume that grounded effects override embodied effects in the vertical dimension, which means a grounded effect should prevail.

In the present study, we aimed at evaluating this hypothesis. We presented nouns referring to objects typically located in lower vs. upper vertical space (e.g., “worms” vs. “birds”) and spatially neutral nouns (e.g., “machines”) in plural form after either a low or high number prime. Put differently, our words were congruent or incongruent with the number cues with respect to two different dimensions. With respect to their semantics, up-words are congruent with high numbers and incongruent with low numbers whereas down-words are congruent with low numbers and incongruent with high numbers. With respect to their grammatical number, up- and down-words are both congruent with high numbers and incongruent with low numbers. Neutral plural nouns, in contrast are only congruent or incongruent with respect to one dimension, namely grammatical number; they are congruent with high number cues and incongruent with low number cues (see **Table 1**). In our study, we were interested

in evaluating the relative impact of congruency on the two dimensions.

Against the above described background, our hypotheses were as follows. According to Fischer and Brugger (2011), the representation of number magnitude is assumed to be grounded in the vertical spatial dimension whereas the representation of syntactic grammatical multitude is assumed to be embodied on a horizontal spatial dimension (Roettger and Domahs, 2015). As such (i) due to their grounding on vertical space, an effect of congruency for numerical cues and word meaning associated with lower and upper space should be observed with faster reaction times for congruent number-word pairs (low numbers/down-words, high numbers/up-words) compared to incongruent pairs (low numbers/up-words, high numbers/down-words, cf. Lachmair et al., 2014). Moreover, this line of argument would suggest (ii) an embodied effect of numerical cues on spatially neutral words due to their plural word form with faster reaction times for high numbers compared to low numbers as shown in Roettger and Domahs (2015). However note, due to the hierarchical superiority of grounded over embodied effects, it is also possible that grounded influences may be processed predominantly by definition. As such, the preference of grounded effects may generally reduce the probability to observe embodied effects such as (ii). Importantly, because of the different nature of the two potential influences (grounded vs. embodied) there should be (iii) no interaction between the two. In other words, the congruency effect in (i) should not be affected by congruency with respect to grammatical number. There is, thus, no reason to expect that the congruency effect between numerical cues and word meaning will differ between up- and down-words. We will refer to this hypothesis as the Grounded-Embodied-Hypothesis in the following.

However, following Berent et al. (2005) both representations of numerical magnitude and syntactic grammatical multitude are grounded. This hypothesis predicts (i) and (ii) as above but without the possibility of (ii) being overridden by (i). Importantly, in contrast to the above discussed Grounded-Embodied-Hypothesis, this hypothesis would predict congruency effects to differ between up- and down-words. In particular, for up words, where congruency with respect to word meaning and congruency with respect to grammatical number fall together, a larger overall congruency effect is to be expected. Contrarily, for down-words incongruence on the two dimensions should result in a smaller overall congruency effect. We will refer to this hypothesis as the Grounded-Grounded-Hypothesis in the following.

TABLE 1 | Congruency of numbers and words according to word meaning or grammatical number. “+” denotes congruency, “-” denotes incongruency and “°” neither congruency nor incongruency.

	word meaning			grammatical number		
	Up	Down	Neutral	Up	Down	Neutral
2, 3	-	+	°	-	-	-
8, 9	+	-	°	+	+	+

MATERIALS AND METHODS

Participants performed a lexical decision task on plural nouns denoting objects that are typically encountered in the upper or lower vertical space, as well as spatially neutral words (e.g., *roofs* vs. *roots* vs. *machines*, respectively). These nouns were preceded by either small (2, 3) or large number cues (8, 9). Please note, the study by Lachmair et al. (2014) investigated priming effects of numbers “1,” “2,” “8,” and “9” on words in singular word form. However, using “1” as a cue might lead to conflicts when processing the plural word form employed in this study. Therefore, the number cue “1” was replaced by the number cue “3,” so that all number cues denoted plurality and would not interfere with our study goals.

Participants

Twenty-two right-handed native speakers of German (17 female; $M_{\text{age}} = 22.64$ years, $SD = 3.17$) took part in this experiment. Experimental testing was in agreement with the guidelines for good scientific practice at the University of Tübingen (Germany). Participants' anonymity was always preserved. All participants gave their written informed consent and received course credit or financial reimbursement of 8 Euros per hour for participation. All participants had normal or corrected-to-normal vision.

Materials and Apparatus

Materials consisted of the numbers 2, 3, 8, and 9, as well as 60 German nouns and 20 pseudo words. Of the 60 nouns, 20 referred to an object that is typically located in upper vertical space, 20 referred to objects that are typically located in lower vertical space and 20 referred to objects denoting a neutral position according to verticality. All nouns were taken from the study by Lachmair et al. (2011), being controlled for frequency, length and for the typical vertical position of their referent (cf. Lachmair et al., 2011). Words and numbers were presented in white against a black background on a 17" CRT monitor. The vertical visual angle varied according to word length between 2.15° and 5.4°. Responses were recorded using a standard QWERTZ keyboard with horizontally aligned response keys. We employed the ‘y’-key for left hand responses and the ‘z’-key for right hand responses.

Procedure and Design

Participants were presented with plural nouns preceded by a one-digit number prime (i.e., 2, 3, 8, or 9). Primes and subsequent nouns were presented in the center of the screen. Participants had to decide whether the presented letter string was a correct German word or not. Each participant started with a short practice block (32 trials) consisting of words of the word-categories UP, DOWN, NEUTRAL and PSEUDOWORDS presented subsequent to numerical cues. Then, in the first half of the experiment, participants had to respond with a left key press to words and a right key press to pseudo-words. With another 32 practice trials the second half of the experiment started in which hand-to-response mapping was reversed. Each trial started with a centered fixation cross (500 ms), followed by a number prime presented for 300 ms. Then the (pseudo)-word appeared and stayed on the screen until a response occurred.

Response times (RTs) were measured as the time from word onset to a key response. Each stimulus was presented eight times (four times in each half), resulting in a total of 640 experimental trials (480 word-trials and 160 pseudo word-trials), subdivided into 8 blocks, separated by self-paced breaks with error information. Each experimental half started with a short practice block. The design was a 2×3 design with the numerical magnitude of the number cues (low, high) and the implicit locational association of words (word category: up, down, neutral) as within-participant factors. Please note, that the locations of the response keys to the left and to the right were not important for the design, because their spatial alignment was horizontal, not vertical and the mapping with pseudo and non-pseudo words was counterbalanced.

RESULTS

All data were analyzed using R (R Development Core Team, 2017). The data of one participant had to be excluded due to an error rate exceeding 20%. Responses to pseudo words were excluded from analyses. A trimming procedure further eliminated responses slower than 200 ms (0.03%), erroneous responses (2.54%), as well as responses for which RT deviated by more than 3 SDs from the individual's mean in the respective condition. This led to an additional loss of 1.91% of the data. The means of the remaining reaction times are displayed in **Figure 1** as a function of word category and number cue magnitude. For investigating our hypotheses, we conducted a 2×3 ANOVA with the within-factors number cue magnitude (low vs. high) and word category (up vs. down vs. neutral).

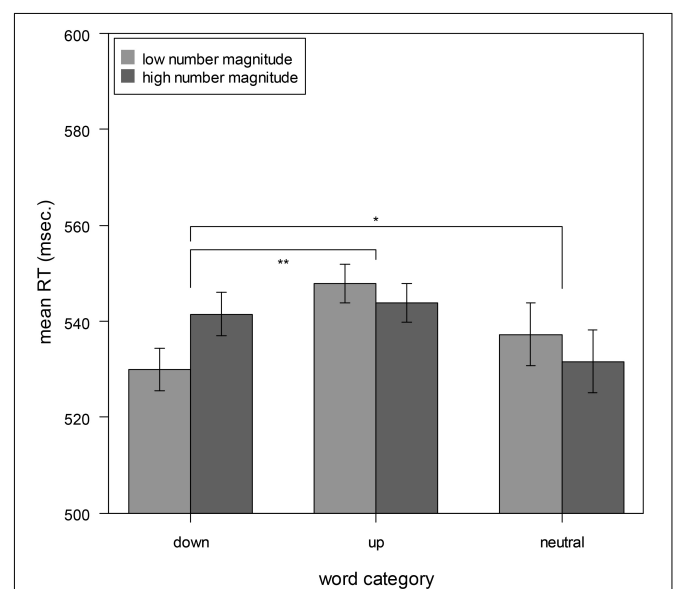


FIGURE 1 | Mean reaction times as a function of implicit locational association of words (up, down, neutral) and numbers (high, low). Error bars represent 95% confidence intervals. * $p < 0.05$; ** $p < 0.01$.

The ANOVA revealed a significant main effect of word category [$F(2,40) = 4.21$, $p = 0.022$, $\eta_p^2 = 0.17$] indicating slower responses to up-words ($RT_{\text{mean}} = 546$ ms, $SD = 136$ ms) compared to down- ($RT_{\text{mean}} = 536$ ms, $SD = 131$ ms) and neutral-words ($RT_{\text{mean}} = 534$ ms, $SD = 126$ ms). Additionally, there was a significant interaction between number cue magnitude and word category [$F(2,40) = 4.40$, $p = 0.019$, $\eta_p^2 = 0.18$]. To break down this 2×3 interaction, we conducted several additional analyses.

First, we excluded the neutral words and conducted a 2 (number cue magnitude: low vs. high) \times 2 (word category: up vs. down) ANOVA which revealed a significant two way interaction [$F(1,20) = 8.35$, $p = 0.009$, $\eta_p^2 = 0.29$]. As can be seen from inspecting the means in **Figure 1**, reaction times were shorter in congruent conditions (i.e., high numbers followed by up-words and low numbers followed by down-words) compared to incongruent conditions (i.e., low numbers followed by up-words and high numbers followed by down-words), and the difference between the congruent and the incongruent condition was numerically larger for down-words than for up-words.

Second, we excluded the up-words and conducted a 2 (number cue magnitude: low vs. high) \times 2 (word category: down vs. neutral) ANOVA. This ANOVA also revealed a significant two-way interaction [$F(1,20) = 5.8$, $p = 0.026$, $\eta_p^2 = 0.23$]. Again, reaction times in the congruent condition were shorter than those in the incongruent condition, with the difference being numerically larger for down words than for neutral words.

Third, we excluded the down-words and conducted a 2 (number cue magnitude: low vs. high) \times 2 (word category: up vs. neutral). This ANOVA did not show a significant interaction effect ($F < 1$, $p = 0.81$).

Finally, evaluating simple effects t -tests revealed for down-words significantly faster RTs when they followed a low number cue ($RT_{\text{mean}} = 530$ ms, $SD = 50$ ms) compared to a high number cue ($RT_{\text{mean}} = 542$ ms, $SD = 57$ ms; $t(20) = 3.16$, $p = 0.005$, $\eta_p^2 = 0.33$). However, for up-words, a t -test indicated no significant advantage of RTs when they were presented following a high number cue ($RT_{\text{mean}} = 544$ ms, $SD = 64$ ms) as compared to a low-number cue ($RT_{\text{mean}} = 548$ ms, $SD = 62$ ms; $t(20) = -1.44$, $p = 0.16$, $\eta_p^2 = 0.09$). A similar finding was obtained for neutral words for which RTs did not differ significantly following a low ($RT_{\text{mean}} = 537$ ms, $SD = 61$ ms) or high number cue ($RT_{\text{mean}} = 532$ ms, $SD = 56$ ms; $t(20) = -0.96$, $p = 0.35$, $\eta_p^2 = 0.04$, see **Figure 1**).

DISCUSSION

Recent research indicated spatial associations for words referring to entities with a typical location in vertical space, as well as for numbers. In the current study, we were interested in the interrelation between these spatial associations. Specific attention was paid to the role played by the magnitude and multitude status of the words. Participants were presented with a numerical cue (low: 2, 3 vs. high: 8, 9) and a subsequent word in plural flexion. These words were nouns that referred to entities typically located in lower or upper vertical space (e.g., roots vs. roofs) or spatially neutral nouns (e.g., “tables”). Considering the idea of a

grounding of numbers and word meanings in vertical space, we evaluated whether the congruency between number magnitude and spatial aspects of word meaning generalizes to plural word forms, and if so how this effect is affected by grammatical number.

Accordingly, two hypotheses were formulated. The Grounded-Embodied-Hypothesis predicts (i) a congruency between numerical cues and word meanings associated with lower and upper space according to their grounding in vertical space, and (ii) an embodied effect of numerical cues on spatially neutral words due to their plural word form. However, according to Fischer and Brugger (2011) the grounded effect of (i) may also override the embodied effect of (ii) causing the latter not to show.

In contrast, the Grounded-Grounded-Hypothesis would predict a more robust influence of numerical cues on spatially neutral words due to their plural word form which should not be overridden by spatial congruency of numbers and up- and down words. In addition, an influence of grammatical number on the congruency effect between number magnitude and spatial cues conveyed by word meaning would be expected. In particular, a larger congruency effect should be observed for up-words than for down words (see above).

Our results substantiated the Grounded-Embodied-Hypothesis: First, we observed a significant interaction between the magnitude of numerical cues and the word meaning of up vs. down words. We observed faster reaction times for congruent number-word pairs (high number/up-word, low number/down-word) compared to incongruent number-word pairs (high number/down-word, low number/up-word). Second, it appears that the difference of reaction times between low and high number cues was more pronounced for down-words than for up-words, which is opposite to what was expected from the Grounded-Grounded-Hypothesis.

Moreover, given that no congruency effect was observed for syntactic grammatical number for the neutral words, one might conclude that a spatial mapping for multitude as suggested by Roettger and Domahs (2015) may not have been sufficiently activated. This claim is further corroborated by additional analyses more closely reflecting analyses and results of Roettger and Domahs (2015) who primarily observed the congruency effect in late processing stages. When we only considered reaction times larger than the median of each participant for neutral words following high or low number cues, this did not reveal any indication of a congruency effect according to grammatical number for neutral words either.

In our view, there exist two possible, not mutually exclusive explanations for this pattern of results, which is in contrast to the study by Roettger and Domahs (2015).

First, we focused on the vertical dimension, while Röttger and Domahs focused on the horizontal one. As we laid out in the introduction, directional spatial-numerical associations in the vertical dimension are assumed to be more grounded, while horizontal ones are thought to be more embodied. Because directional associations of numbers and space are related to reading direction (which is an embodied experience), it is conceivable that grammatical number as a language attribute

may also be more prone to embodied influences. However, as embodied influences are weaker in the vertical condition, these might not have been sufficient to automatically activate a directional association of multitude (grammatical number) with number magnitude.

Second, we presented participants only with plural nouns and not with singular and plural words in one experiment. This might have decreased the saliency of grammatical number in contrast to Roettger and Domahs (2015) in two ways. First, because there is no variation of singular and plural, neither in the grammatical forms of the nouns, nor in the grammatical number associated with the digits (i.e., also only plural because 1 was excluded), grammatical number may not have been salient enough to influence results significantly. Second, and maybe even more importantly, multitude and grammatical number was not task-relevant. This may have also reduced saliency. However, based on the current data, we can at least infer that the activation of grammatical number may not be as automatic as suggested, for example, by Berent et al. (2005). Clearly, this issue deserves further investigation in the future.

In contrast to the lack of effects for multitude, our data suggest that co-activation of spatial attributes of number magnitude and the implicit down- and up-ward associations of up and down words are due to automatic processes resulting from the Groundedness. In turn, this led to the obtained congruency effect. Thus, considering the proposition of Fischer and Brugger (2011, see also Myachykov et al., 2014), our results substantiate the hypothesis that associations between implicit down- and up-ward attributes of word meaning and number magnitude are spatially grounded. Note, however, that in contrast to multitude, both number magnitude (small 2, 3 vs. large 8, 9) and word meaning (down-words, up-words, neutral words) was varied. This was not the case for multitude (only plural words were used), which might have played a role in the pattern of results obtained, even though both number magnitude and word meaning were also not task-relevant in our lexical decision task.

Interestingly, these data are also consistent with the notion that linguistic influences on number processing seem to occur on different representational levels. In their recent taxonomy of linguistic influences on number processing, Dowker and Nuerk (2016) differentiate between several linguistic levels at which number processing may be influenced. For the current study,

influences on the syntactic and the semantic level are most relevant because the association between numerical magnitude (low/high) and word meaning (e.g., roots/roofs) is driven by the semantics of the words. In contrast, the association of number magnitude (low/high) and grammatical number (singular/plural) refers to the syntactic attribute of grammatical number. The observed result pattern suggests the association of number magnitude and word meaning to be grounded according to the framework of Fischer and Brugger (2011). This may have prevented the observation of an association of number magnitude and syntactic grammatical number, which is considered to be embodied in the horizontal dimension (cf. Roettger and Domahs, 2015). As such, this implies that semantic and syntactic linguistic influences on number processing may not interact on the same representational level. Instead, associations at the level of the meaning of words (i.e., up- vs. down-words) and numbers (i.e., the numerosity they reflect) seem to be more prominent as compared to associations across semantic (i.e., the numerosity they reflect) and syntactic (grammatical number) levels.

In summary, the present study showed a spatial congruency between low and high number magnitude cues (e.g., 2 vs. 8) and words referring to objects up or down in the world presented in plural word form. No influence of grammatical number on spatially neutral words or on the spatial congruency effect was found. Thus, together with the results of the study by Lachmair et al. (2014) this supports the view of a grounded spatial congruency between numbers and word meaning regardless of the syntactical word form. Future research is needed to substantiate this claim and to investigate (i) whether it is a general pattern that associations are most prominent when levels of linguistic and numerical processing match or (ii) whether certain (situated) experimental conditions moderate or mediate the differences observed between associations of magnitude or multitude of numbers and word meaning.

AUTHOR CONTRIBUTIONS

ML, BK and H-CN: conception. ML: analysis and drafting. ML, SRF, H-CN, BK and KM: interpretation, approval, and agreement. BK, SRF, H-CN and KM: revising.

REFERENCES

- Anobile, G., Cicchini, G. M., and Burr, D. C. (2016). Number as a primary perceptual attribute: a review. *Perception* 45, 5–31. doi: 10.1177/0301006615602599
- Barsalou, L. W. (2012). “The human conceptual system,” in *The Cambridge Handbook of Psycholinguistics*, eds M. J. Spivey, K. McRae, and M. Joanisse (New York, NY: Cambridge University Press), 239–258. doi: 10.1017/CBO9781139029377.017
- Berent, I., Pinker, S., Tzelgov, J., Bibi, U., and Goldfarb, L. (2005). Computation of semantic number from morphological information. *J. Mem. Lang.* 53, 342–358. doi: 10.1016/j.jml.2005.05.002
- Cipora, K., Patro, K., and Nuerk, H. C. (2015). Are spatial-numerical associations a cornerstone for arithmetic learning? The lack of genuine correlations suggests no. *Mind Brain Educ.* 9, 190–206. doi: 10.1016/j.dcn.2017.06.001
- Dehaene, S. (2011). *The Number Sense: How the Mind Creates Mathematics*. New York, NY: Oxford University Press.
- Dehaene, S., Bossini, S., and Giraux, P. (1993). The mental representation of parity and number magnitude. *J. Exp. Psychol. General* 122, 371. doi: 10.1037/0096-3445.122.3.371
- Dowker, A., and Nuerk, H. C. (2016). Linguistic influences on mathematics. *Front. Psychol.* 7:1035. doi: 10.3389/fpsyg.2016.01035
- Downing, P. A. (1996). *Numerical Classifier Systems: The Case of Japanese*, Vol. 4. Philadelphia, PA: John Benjamins Publishing. doi: 10.1075/sidag.4
- Dudschig, C., Souman, J., Lachmair, M., de la Vega, I., and Kaup, B. (2013). Reading “sun” and looking up: the influence of language on saccadic eye movements in the vertical dimension. *PLoS One* 8:e56872. doi: 10.1371/journal.pone.0056872
- Fischer, M. H., and Brugger, P. (2011). When digits help digits: spatial-numerical associations point to finger counting as prime example of embodied cognition. *Front. Psychol.* 2:260. doi: 10.3389/fpsyg.2011.00260

- Fischer, M. H., and Shaki, S. (2014). Spatial associations in numerical cognition—From single digits to arithmetic. *Q. J. Exp. Psychol.* 67, 1461–1483. doi: 10.1080/17470218.2014.927515
- Fodor, J. A. (1975). *The Language of Thought*, Vol. 5. Cambridge, MA: Harvard University Press.
- Holmes, K. J., and Lourenco, S. F. (2012). Orienting numbers in mental space: horizontal organization trumps vertical. *Q. J. Exp. Psychol.* 65, 1044–1051. doi: 10.1080/17470218.2012.685079
- Lachmair, M., Cress, U., Fissler, T., Kurek, S., Leininger, J., and Nuerk, H. C. (2017). Music-space associations are grounded, embodied and situated: examination of cello experts and non-musicians in a standard tone discrimination task. *Psychol. Res.* doi: 10.1007/s00426-017-0898-y [Epub ahead of print].
- Lachmair, M., Dudschig, C., De Filippis, M., de la Vega, I., and Kaup, B. (2011). Root versus roof: automatic activation of location information during word processing. *Psychon. Bull. Rev.* 18, 1180–1188. doi: 10.3758/s13423-011-0158-x
- Lachmair, M., Dudschig, C., de la Vega, I., and Kaup, B. (2014). Relating numeric cognition and language processing: do numbers and words share a common representational platform? *Acta Psychol.* 148, 107–114. doi: 10.1016/j.actpsy.2013.12.004
- Lakoff, G., and Johnson, M. (2008). *Metaphors We Live By*. Chicago, IL: University of Chicago press.
- Levinson, S. C. (2003). *Space in Language and Cognition: Explorations in Cognitive Diversity*, Vol. 5. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511613609
- Myachykov, A., Scheepers, C., Fischer, M. H., and Kessler, K. (2014). TEST: a tropic, embodied, and situated theory of cognition. *Top. Cogn. Sci.* 6, 442–460. doi: 10.1111/tops.12024
- Overmann, K. A. (2015). Numerosity structures the expression of quantity in lexical numbers and grammatical number. *Curr. Anthropol.* 56, 638–653. doi: 10.1086/683092
- R Development Core Team (2017). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Restle, F. (1970). Speed of adding and comparing numbers. *J. Exp. Psychol.* 83(2Pt 1), 274–278. doi: 10.1186/1471-2334-11-269
- Roettger, T. B., and Domahs, F. (2015). Grammatical number elicits SNARC and MARC effects as a function of task demands. *Q. J. Exp. Psychol.* 68, 1231–1248. doi: 10.1080/17470218.2014.979843
- Sarnecka, B. W. (2014). On the relation between grammatical number and cardinal numbers in development. *Front. Psychol.* 5:1132. doi: 10.3389/fpsyg.2014.01132
- Schwarz, W., and Keus, I. M. (2004). Moving the eyes along the mental number line: comparing SNARC effects with saccadic and manual responses. *Attent. Percept. Psychophys.* 66, 651–664. doi: 10.3758/BF03194909
- Thornton, T., Loetscher, T., Yates, M. J., and Nicholls, M. E. (2013). The highs and lows of the interaction between word meaning and space. *J. Exp. Psychol. Hum. Percept. Perform.* 39, 964–973. doi: 10.1037/a0030467
- Tulving, E. (1972). Episodic and semantic memory. *Organ. Mem.* 1, 381–403.
- Winter, B., Matlock, T., Shaki, S., and Fischer, M. H. (2015). Mental number space in three dimensions. *Neurosci. Biobehav. Rev.* 57, 209–219. doi: 10.1016/j.neubiorev.2015.09.005

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling Editor declared a past co-authorship with one of the authors H-CN.

Copyright © 2018 Lachmair, Ruiz Fernández, Moeller, Nuerk and Kaup. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Numerical Affordance Influences Action Execution: A Kinematic Study of Finger Movement

Rosa Rugani^{1*}, Sonia Betti¹ and Luisa Sartori^{1,2*}

¹ Department of General Psychology, University of Padua, Padua, Italy, ² Padova Neuroscience Center, University of Padua, Padua, Italy

OPEN ACCESS

Edited by:

Hans-Christoph Nuerk,
Universität Tübingen, Germany

Reviewed by:

Samuel Shaki,
Ariel University, Israel
Thomas J. Faulkenberry,
Tarleton State University,
United States

*Correspondence:

Rosa Rugani
rosa.rugani@unipd.it
Luisa Sartori
luisa.sartori@unipd.it

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 13 December 2017

Accepted: 16 April 2018

Published: 01 May 2018

Citation:

Rugani R, Betti S and Sartori L
(2018) Numerical Affordance
Influences Action Execution:
A Kinematic Study of Finger
Movement. *Front. Psychol.* 9:637.
doi: 10.3389/fpsyg.2018.00637

Humans represent symbolic numbers as oriented from left to right: the mental number line (MNL). Up to now, scientific studies have mainly investigated the MNL by means of response times. However, the existing knowledge on the MNL can be advantaged by studies on motor patterns while responding to a number. Cognitive representations, in fact, cannot be fully understood without considering their impact on actions. Here we investigated whether a motor response can be influenced by number processing. Participants seated in front of a little soccer goal. On each trial they were visually presented with a numerical (2, 5, 8) or a non-numerical (\$) stimulus. They were instructed to kick a small ball with their right index toward a frontal soccer goal as soon as a stimulus appeared on a screen. However, they had to refrain from kicking when number five was presented (no-go signal). Our main finding is that performing a kicking action after observation of the larger digit proved to be more efficient: the trajectory path was shorter and lower on the surface, velocity peak was anticipated. The smaller number, instead, specifically altered the temporal and spatial aspects of trajectories, leading to more prolonged left deviations. This is the first experimental demonstration that the reaching component of a movement is influenced by number magnitude. Since this paradigm does not require any verbal skill and non-symbolic stimuli (array of dots) can be used, it could be fruitfully adopted to evaluate number abilities in children and even preschoolers. Notably, this is a self-motivating and engaging task, which might help children to get involved and to reduce potential arousal connected to institutional paper-and-pencil examinations.

Keywords: mental number line, spatial-numerical association, kinematics, reaching, action execution, finger movement, numerical cognition

INTRODUCTION

The propensity to spatially represent environmental information is a core characteristic of human cognitive system (Gevers et al., 2003). Numbers are coded into space along a left-right oriented continuum (Fias and Fischer, 2005; Buetti and Walsh, 2009; Dehaene, 2011). The seminal insight of such a spatial-numerical association goes back to 1880, when Galton (1880) firstly proposed that humans describe and think numbers as increasingly oriented from left to right along a mental number line (MNL), where small numbers are located on the left and large numbers on the right side of space. The first scientific demonstration of this spatial representation of number has been

reported more than 100 years later, when Dehaene et al. (1993) discovered that humans respond faster to smaller numbers on the left space and to larger numbers on the right space; the Spatial Numerical Association of Response Codes (SNARC) effect. A large body of literature supports this effect. Humans show a left bias when indicating the center of a string composed of repeated “1”, and a right bias when it is composed of “9” (Fischer, 2001). This indicates that an automatic activation of the left or right space automatically occurs during number processing: the elaboration of small numbers pre-activates the left space and the elaboration of large numerical magnitudes pre-activates the right space.

Complementary results have been obtained in a pseudo-random number generation task. Loetscher et al. (2008) asked participants to report random numbers in the 1–30 numerical range. Participants were systematically influenced by the side (left or right) their head was turned. When they faced toward their left, they produced comparatively more small numbers with respect to when they faced toward their right (Loetscher et al., 2008; see also Winter et al., 2015 for similar results along both near/far space and vertical dimensions, and Hartmann et al., 2012 for a whole body condition). Passive observation of leftward or downward gaze similarly induced participants to generate smaller than large numbers, compared to observing color changes or rightward gaze (Grade et al., 2013). Such biases are explained by a shifting of the attention compatible with the MNL, which facilitates the accessibility to small numbers turning the left and to large numbers turning the right. More recently, the effect reported by Loetscher et al. (2008) has been replicated in a condition of lateral arm turns. The effect was present when two congruent body’s movements were required (e.g., right-turns of both arm and head), but it disappeared whenever the two movements were incongruent (e.g., left-turns of arm and right-turns of head). This reveals that the spatial bias induced by the two sensorimotor locations on numerical processing can annihilate each other (Cheng et al., 2015). All together, these findings show that numbers and motor actions influence each other. This interaction is not limited to laboratory experiences but it emerges also in everyday activities. Numerical magnitude influences directional decisions while walking. In a recent study, healthy adults were required to stand and to produce random numbers as they made lateral turns. Lateral turn decisions could be predicted by the magnitude of random numbers produced before the turn: participants turned left more often when they had just produced small numbers, *vice-versa* they turned right more often when they had just produced large numbers (Shaki and Fischer, 2014).

Since cognitive representations of perceptual and semantic information are fully understood only when considering their impact on actions (Gallese and Lakoff, 2005), the existing knowledge on the MNL should be extended to studies that analyze motor actions while responding to a number (see for example Girelli et al., 2016). An emerging literature of hand-tracking and computer-mouse tracking nicely depict how motor actions can be better understood while performing number related tasks (Song and Nakayama, 2008;

Santens et al., 2011; Dotan and Dehaene, 2013; Faulkenberry, 2014; Faulkenberry et al., 2016, 2017).

From this fascinating perspective, adopting kinematic measures is a state-of-art methodology in order to provide a fine-tuned analysis of movement, a large range of degrees of freedom and a highly sensitive investigation. In fact, a mounting number of studies are now using 3-D motion capture and detailed kinematic analyses to measure behavior and to deeply examine questions relating to cognitive processing in naturalistic protocols (for reviews, see Castiello, 2005; Krishnan-Barman et al., 2017). A growing number of studies on prehension movements is proving that semantic information related to magnitude can indeed influence movement kinematics. In particular, it has been shown that grip aperture varies according to the dimension indicated by a label put on a target object: it is larger for the large-labeled object and smaller for the small-labeled object (Gentilucci et al., 2000, 2012; Andres et al., 2008a,b; Namdar et al., 2014). Precision grip movements are faster in response to small numbers and power grips are faster in response to large numbers (Lindemann et al., 2007). These studies clearly show the influence of numbers on motor patterns. However, this effect could also reflect a highly overlearned motor association between magnitude labels (e.g., small, medium, large) and manual responses (e.g., grasping a small or large glass of coke, a 0.5 kg or a 1 kg flour packet). These frequent experiences, though allowing to perform very efficient actions in everyday life (Schwarz and Keus, 2004), could bias to perform smaller grasping actions in relation to smaller digits and vice versa (for review, see Rugani and Sartori, 2016). Notably, two components characterize prehension movements (Jeannerod, 1981; Jakobson and Goodale, 1991; Chieffi and Gentilucci, 1993; Castiello, 1996; Smeets and Brenner, 1999). The reaching component extracts information regarding the object’s spatial location and activates those muscles relevant to approach it. The grasping component extracts information on the object’s intrinsic properties such as size and shape. The open question is whether number processing influences only the grip component or the preceding reaching movement as well (grasp and transport components). To pursue this question in an unbiased way, we recently adopted a new and not-overlearned paradigm (Rugani et al., 2017; see also Betti et al., 2015 for a previous application of this paradigm). We specifically combined a “free response” task with the kinematic analysis of a finger movement and we provided the first demonstration that numerical processing affect not only the grasping, but also the reaching component of movements. This finding particularly depicts the novelty of our approach: instead of measuring the grasping component – which might be affected by previous experience – we adopted a culturally unbiased index (i.e., the transport component). Participants were seated in front of two little soccer goals, one on their left and one on their right side, and they were instructed to kick a small ball with their right index toward the goal indicated by an arrow on the monitor. In a few crucial trials participants were presented also with a small (2) or a large (8) number, and they were allowed to choose the kicking direction. Participants performed more left responses with the small number and more right responses with the large number. The whole kicking movement was then

segmented in two temporal phases (i.e., Kick Preparation and Kick Finalization) in order to make a fine-grained analysis of action execution timing. Results showed that in responding to small numbers toward the left and to large numbers toward the right, participants were faster to finalize the action. Moreover, the small number specifically altered the temporal and spatial aspects of left kick's trajectories, whereas the large number specifically modified right kick's trajectories. However, a limit of that study is that data concerning the two different movements (i.e., left and right kicks) had to be considered separately due to mechanical and anatomical differences (i.e., the degrees of freedom of the right index finger in relation with the anatomy of the right hand). Here, we adopted a unique action – a straight kick – for all the experimental conditions to avoid any anatomical bias. This means that we expected all the kicks to differ in terms of temporal features of trajectory path, rather than spatial features (i.e., a general leftward deviation was expected across conditions given the degrees of freedom of the right index during the kicking).

Extensive literature on reach-to-grasp consistently showed a general anticipation in hand kinematics when a target object has to be carefully approached (e.g., with the intention to pour vs. to place it, see Schuboe et al., 2008; or with the intention to throw it vs. to lift it, see Armbrüster and Spijkers, 2006). Moreover, it is known that object weight influences motor planning and control of reach-to-grasp actions as to guarantee a stable final grip placement on the object (Weir et al., 1991; Brouwer et al., 2006; Eastough and Edwards, 2007). In particular, Ansuini et al. (2016) recently found that peak velocity between 10 and 40% of normalized movement time was greater when reaching an heavy than a light object. Interestingly, object weight can also influence simply pantomimed reach-to-grasp actions, thus reflecting a link between cognitive representations of the weight and distinctive features of a motor act.

Since the standard parameters utilized for characterizing the reaching component are essentially trajectory and velocity, here we expect that a functional connection between numerical cognition and action planning will translate into different spatial and temporal patterns across conditions. Since numerical priming has two features: (i) spatiality (small numbers are associated with left space and large numbers with right space) and (ii) weight (small numbers are associated with light objects and large numbers with heavy objects), these two features should

jointly influence hand movement kinematics. In particular, we predict that the smaller number will influence the temporal aspects of left trajectory deviations, in line with our previous study (Rugani et al., 2017). Whereas the larger number will produce a more direct route, as indexed by lower and shorter trajectory path, and anticipated peak velocity. This innovative approach combining number presentation with action execution will allow us to measure number-related information transmitted by the hand movements over time.

MATERIALS AND METHODS

Participants

Twenty-three students (10 males and 13 females, mean age = 22.74 years, SD = 0.75) took part in the experiment. A statistical power analysis for sample size estimation was previously performed (GPower 3.1), based on data from a published study (Rugani et al., 2017). The mean effect size (ES) of paired *t*-test in that study (0.65) was considered to be large/medium according to Cohen's (1988) criteria. Here, since we planned to use a repeated-measure ANOVA, for sample size estimation we inserted these values: $\eta^2 = 0.20$; $\alpha = 0.01$; $1 - \beta = 0.99$; number of measures = 3; groups = 1; supposed correlation among measures = 0.45. The projected sample size needed with this effect size is $N = 23$ for within group comparisons. All participants were right handed, had normal or corrected-to-normal vision, and were naive about the purpose of the experiment. Participants gave their written consent before the experiment. The experimental procedures were approved by the Ethics Committee of the University of Padova and were carried out in accordance with the principles of the 1964 Declaration of Helsinki (Sixth revision, 2008).

Stimuli

Stimuli consisted in three symbolic numbers: a small digit (2), an intermediate digit (5), and a large digit (8), plus a symbolic character semantically associated with numbers, though not a number in itself (\$, see **Figure 1**). This character was specifically selected on the basis of its symmetry, in order to avoid any indication of direction (as compared to #, for example, which is slightly tilted to the right). The stimulus 5 was adopted as a no-go signal to ensure that reaching movements were not

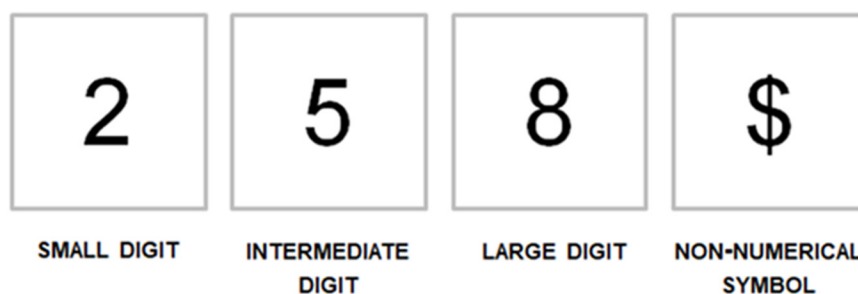


FIGURE 1 | Stimuli adopted in the experiment. A small digit (2), an intermediate digit (5; no-go signal), a large digit (8), and a non-numerical symbol (\$).

initiated before the number was processed. Hereafter, stimuli will be referred to as S2, S5, S8, and S\$. Digits were in Arial font, black color and 160 size. On each trial, a black fixation cross (7.5 cm by 7.5 cm, in Arial font, black color) appeared on the screen before stimulus presentation.

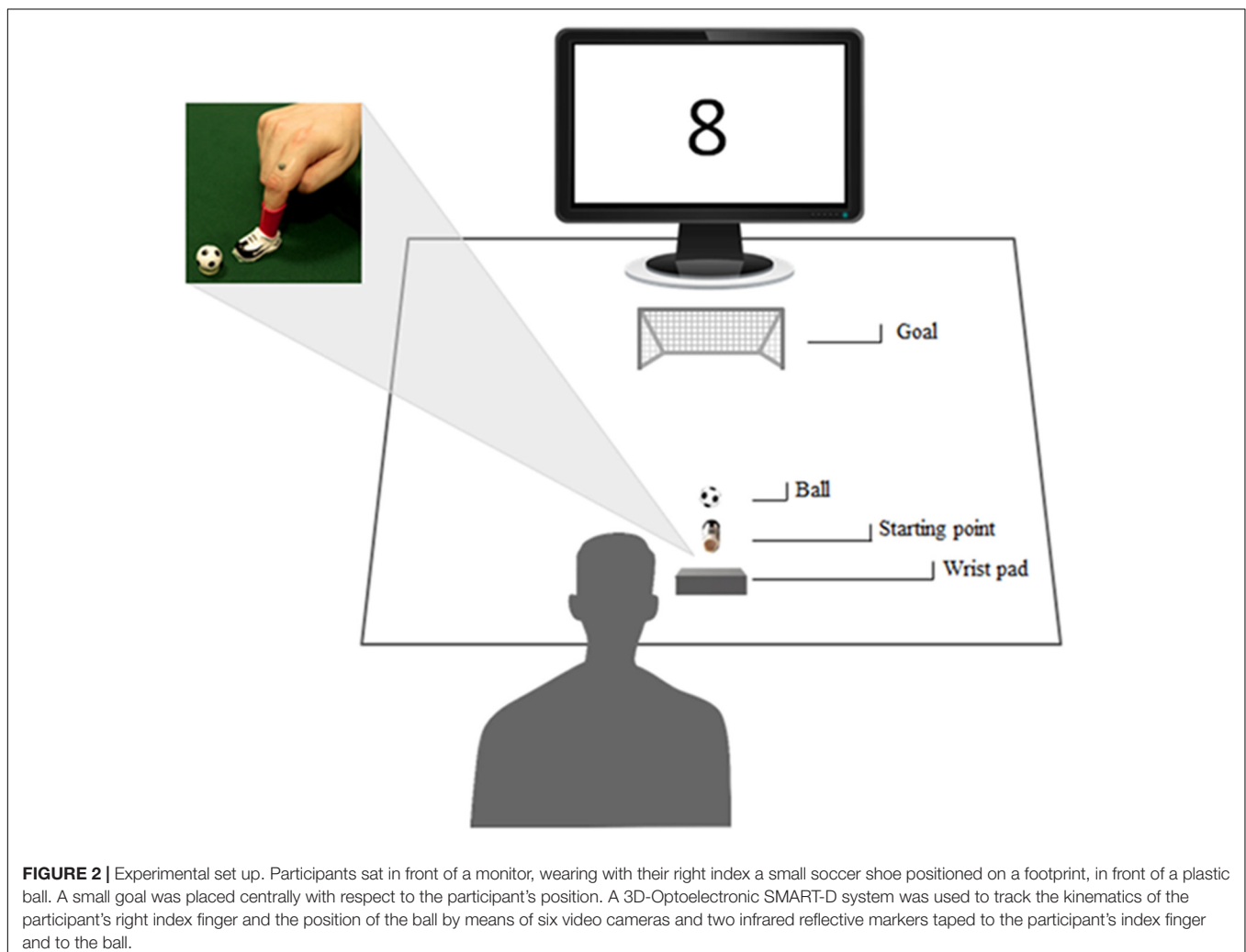
Apparatus and Procedure

Participants sat on a chair in front of a table (90 cm × 90 cm) with the left hand resting on their left leg and the right hand located in the designated start position. The experimental apparatus consisted in a green velvet surface (93.5 cm × 74 cm). Participants' right index was introduced in the plastic sock (4.5 cm high, 2.5 cm diameter) of a small plastic soccer shoe (3 cm long, 1.5 cm wide; for a schematic representation of the apparatus see **Figure 2**). At the beginning of each trial, participants were instructed to position the shoe on a footprint (3 cm long, 1.5 cm wide) painted on the velvet cloth. A plastic ball (2.3 cm of diameter) was positioned on a plastic ring (1.5 cm diameter) located at 1 cm away from the footprint. In the start position, participants rested their right wrist on a pillow (16 cm long, 11 cm wide and 6.5 cm high), which was shaped to guarantee a

comfortable and repeatable posture of the hand, allowing them to effortlessly kick the ball. A small soccer goal (18 cm long 16 cm high) was located 50 cm away from the footprint. A 24" monitor (resolution 1920 × 1080 pixels, refresh frequency 120 Hz) set at eye level (the eye–screen distance was 80 cm) was used to present the experimental stimuli. Participants underwent two sessions (i.e., Training and Testing) and were instructed to kick the ball toward the soccer goal following stimulus presentation, at their own pace. No instruction was given concerning the speed of movement. A black fixation cross appeared for 100 ms and was replaced with a stimulus after 1000 ms. During the Training session S\$ was presented for 15 trials. During the Testing session participants kicked the ball upon random presentation of either a symbolic number or a symbolic character: S2, S8, and S\$ were shown 10 times each. Whenever S5 was presented ($n = 10$ trials), participants were required to refrain from kicking the ball.

Kinematic Recording

A 3D-Optoelectronic SMART-D system (Bioengineering Technology and Systems, B|T|S) was used to track the kinematics of the participant's right index. One light-weight



infrared reflective marker (0.25 mm in diameter; B|T|S) was taped on the index finger's proximal phalange to measure the kicking movement (see **Figure 2**). A second marker was located on the ball to compute the midline virtually connecting index finger and target object and to segment the whole movement in a Pre-Contact and a Post-Contact Phase. Six infrared video cameras (sampling rate 140 Hz) detecting the markers' positions in a 3-D space were placed in a semicircle at a distance of 1–1.2 meters from the table. Each camera position, roll angle, zoom, focus, threshold and brightness were calibrated and adjusted to optimize data collection before each experimental session. For the dynamic calibration, a three-marker wand was moved throughout the workspace of interest for 60 s. The measurements were made along the three Cartesian axes [i.e., x (left–right), y (up–down), and z (anterior–posterior)]. The spatial resolution of the recording system was 0.3 mm over the field of view. The standard deviation of the reconstruction error was 0.2 mm for the x , y , and z axes.

Data Processing

Following kinematic data collection, the SMART-D Tracker software package (B|T|S) was used to provide a 3-D reconstruction of the marker positions of each trial as a function of time. The data were then filtered using a finite impulse response linear filter (transition band = 1 Hz, sharpening variable = 2, cut-off frequency = 10 Hz; D'Amico and Ferrigno, 1990, 1992). Movement onset was defined as the time at which the tangential velocity of the finger marker crossed a threshold (5 mm/s) and remained above it for longer than 500 ms. End of movement was defined as the time at which the tangential velocity of the finger marker dropped below the threshold (5 mm/s) after the ball was kicked. The following kinematic parameters were extracted for each individual movement using a custom Protocol run in Matlab, 2014b (The Math Works, Natick, MA, United States):

Movement Time: the time interval between movement onset and end of movement (ms);

Trajectory Path: the length of the index trajectory (mm);

Maximum Trajectory Height: the maximum height of the index trajectory on the y -axis (mm);

Contact Time: the time at which the tangential velocity of the ball crossed a threshold (2 mm/s) and remained above it for longer than 500 ms;

Time to Maximum Velocity: the time at which index velocity was maximum, with respect to movement onset (ms);

Time to Maximum Trajectory Height: the time at which index trajectory was higher, with respect to movement onset (ms);

Time to Maximum Trajectory Deviation: the time at which index trajectory reached the maximum perpendicular deviation from the virtual line linking the starting position with the target object, with respect to movement onset (ms).

The temporal peaks were then normalized with respect to movement time, so that individual speed differences were accounted for:

Contact Time (%): the percentage of movement time at which the tangential velocity of the ball crossed a threshold (2 mm/s) and remained above it for longer than 500 ms;

Time of Maximum Velocity (%): the percentage of movement time at which the index trajectory was at maximum velocity (%);

Time to Maximum Trajectory Height (%): the percentage of movement time at which the index trajectory reached its higher peak (%).

Time to Maximum Trajectory Deviation (%): the percentage of movement time at which index trajectory reached the maximum deviation from the midline (%).

For each participant and kinematic index, we calculated means and relative standard deviations for each type of stimulus (S2, S8, and S\$).

Data Analysis

The mean values for each parameter of interest were determined for each participant and entered into separate repeated-measures ANOVAs with Stimulus (S2, S8, and S\$) as within-subjects factor.

Preliminary analyses were conducted to check for normality, sphericity (Mauchly test), univariate and multivariate outliers, with no violations noted. For the ANOVA the alpha level of p was set <0.01 , in accordance with our power analysis. Main effects were used to explore the means of interest (*post hoc t-test*) and Bonferroni correction was applied (alpha level of $p < 0.05$) to prevent Type-1 errors. Statistical analyses were performed with SPSS 23 (SPSS Inc., Chicago, IL, United States) software.

RESULTS

All the means, medians and standard errors are summarized in **Table 1**.

Movement Time (ms): The ANOVA performed on MT revealed a non-significant effect of Stimulus [$F_{(2,44)} = 3.48$, $p = 0.04$, $\eta_p^2 = 0.14$].

Trajectory Path (mm): The ANOVA performed on the length of the index trajectory revealed a significant effect of Stimulus [$F_{(2,44)} = 4.75$, $p = 0.01$, $\eta_p^2 = 0.18$]. Observing S8 led to a shorter trajectory with respect to observing S2 ($p = 0.01$). This effect was significant also for S8 compared to S\$ ($p = 0.02$).

Maximum Trajectory Height (mm): The ANOVA performed on the maximum height of the index trajectory revealed a significant effect of Stimulus [$F_{(2,44)} = 323.98$, $p < 0.001$, $\eta_p^2 = 0.94$]. Observing S8 led to a lower trajectory with respect to observing S2 ($p < 0.001$). This effect was significant also for S8 compared to S\$ ($p < 0.001$).

Contact Time (%): The ANOVA performed on CT revealed a non-significant effect of Stimulus [$F_{(2,44)} = 0.55$, $p = 0.58$, $\eta_p^2 = 0.02$].

Time to Maximum Velocity (%): The ANOVA performed on the time at which index velocity was maximum revealed a significant effect of Stimulus [$F_{(2,44)} = 13.35$, $p < 0.001$, $\eta_p^2 = 0.38$]. Observing S8 led to an earlier peak with respect to observing S2 ($p < 0.001$). This effect was significant also for S8 compared to S\$ ($p < 0.001$).

TABLE 1 | Statistically significant key kinematic parameters (mean, standard errors, and median per condition) across stimuli.

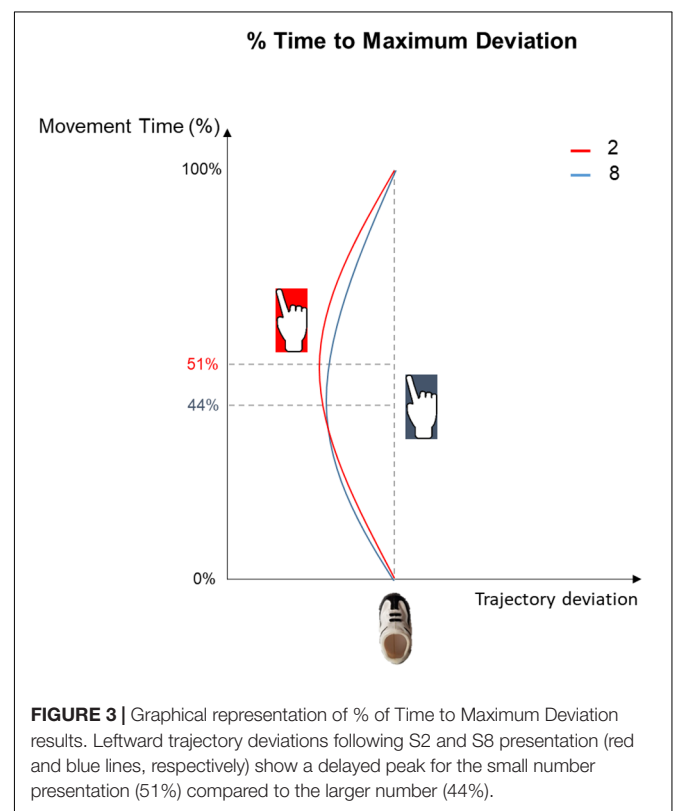
	S2	S8	\$
<i>Movement time (ms)</i>	370.30 (\pm 20.28)	341.12 (\pm 20.11)	359.45 (\pm 18.39)
	376.67	351.52	358.33
<i>Trajectory path (mm)</i>	38.58 (\pm 4.16)	34.95 (\pm 3.84)	37.63 (\pm 4.16)
	34.21	31.41	34.25
<i>Maximum trajectory height (mm)</i>	95.96 (\pm 2.11)	87.07 (\pm 1.85)	95.61 (\pm 1.91)
	94.76	86.97	94.88
<i>Time to maximum velocity (%)</i>	65 (\pm 03)	56 (\pm 03)	64 (\pm 03)
	66	59	64
<i>Time to maximum trajectory height (%)</i>	84 (\pm 04)	77 (\pm 04)	84 (\pm 04)
	94	85	95
<i>Time to maximum trajectory deviation (%)</i>	51 (\pm 03)	44 (\pm 02)	48 (\pm 02)
	52	42	52

Time to Maximum Trajectory Height (%): The ANOVA performed on the time at which index trajectory was higher revealed a significant effect of Stimulus [$F_{(2,44)} = 9.07$, $p < 0.001$, $\eta_p^2 = 0.29$]. Observing S8 led to an earlier peak with respect to observing S2 ($p < 0.001$). This effect was significant also for S8 compared to S\$ ($p = 0.02$).

Time to Maximum Trajectory Deviation (%): The ANOVA performed on the time at which index trajectory reached the maximum deviation from the midline revealed a significant effect of Stimulus [$F_{(2,44)} = 9.07$, $p < 0.001$, $\eta_p^2 = 0.29$]. Observing S2 led to a delayed leftward deviation with respect to observing S8 ($p < 0.001$; see **Figure 3**).

DISCUSSION

The aim of this study was to determine whether number processing affects the performance of executed movements. Participants were asked to perform a kicking action with their right hand after observing a small/large digit (2, 8) or a symbolic character (\$). Our main finding is that although executed actions were exactly the same across conditions, a decrease in Trajectory Path, Trajectory Height, Time to Maximum Velocity and Time to Maximum Trajectory Height occurred for the large compared to the small digit. Our results are in line with previous studies demonstrating a general anticipation when an object is approached more carefully (e.g., Armbrüster and Spijkers, 2006; Schuboe et al., 2008) and an early velocity peak when reaching a heavy than a light object (Ansuini et al., 2016). In our study, performing a finger kicking action after observation of a large digit was indeed highly efficient: the trajectory path was shorter and lower on the surface and the velocity peak was anticipated. In particular, we found an anticipation of the Time of Maximum Velocity ranging from S8 (56%) to S\$ (64%) and S2 (65%), despite the executed movement was the same. Since a statistically significant effect on Time to Maximum Velocity was specifically connected to the observation of S8, this might suggest the activation of an association between larger numbers and weight. By combining knowledge regarding numerical magnitude and weight dynamics,



the motor system might be able to adjust kick kinematics accordingly.

A crucial data arising from the present data is the temporal aspect of trajectory deviations. Given the very short distance between footprint and ball (1 cm) and the constrained end-goal (i.e., straight kick), no effect was expected in terms of trajectory deviations before contact. However, a longer tilt leftwards for the S2 condition during the post-contact phase seems to indicate that participants were aiming towards the left following small number presentation compared to large number presentation. Similar results were obtained with an index finger pointing task (Fischer, 2003). Participants were faster when pointing

leftward after a small digit presentation and rightward after a large digit presentation (Fischer, 2003). Subsequent studies in adults (Ishihara et al., 2006) and in 7 year-old children (Möhrling et al., 2017) revealed that this bias could be explained by a contamination of motor preparation by a direct activation of number magnitude with the congruent spatial location.

Our data show for the first time that the control mechanisms underlying reaching formation are affected by number processing beyond the – already demonstrated – grasping component. Since the effect of numerical magnitude on grip aperture kinematics (Gentilucci et al., 2000, 2012; Andres et al., 2004, 2008a,b; Lindemann et al., 2007; Moretto and di Pellegrino, 2008; Chiou et al., 2012; Namdar et al., 2014) and object affordances (Badets et al., 2007; Chiou et al., 2009) is well known, the present data significantly extend previous literature. The impact of numerical magnitude on both reaching and grasping kinematics would corroborate the theory that representations of number and actions share common codes within a magnitude representation's system (Lindemann et al., 2007).

The ATOM Theory and Numerical Affordance

From a neuropsychological viewpoint, the modulation of numerical cognition on action control could be explained by the “A Theory of Magnitude”, or “ATOM theory” (Walsh, 2003; see also Buetti and Walsh, 2009 for an updated proposal). The ATOM theory postulates that the intra-parietal sulcus (IPS) serves as the cortical center for time, space, and numbers estimation. The IPS would be equipped with an analog system that constantly computes magnitudes for action execution (Buetti and Walsh, 2009). It is therefore plausible that beyond object affordances related to the physical features of an object (Gibson, 1979; Tucker and Ellis, 1998), a “numerical affordance” might link objects' extension and numerosness to specific motor dynamics. Here we specifically demonstrated that two features related to numbers (spatiality and weight) are interrelated and affect movement kinematics. From this perspective, the ATOM theory may explain the interaction between numerical information and non-numerical magnitude, such as time and space, and especially how numbers prompt spatially oriented actions (SNARC effect).

In neural terms, reaching and grasping components are mediated by two separate anatomical pathways (for review see Filimon, 2010). Grasping is organized by a lateral parieto-frontal circuit and reaching by a more medial parieto-frontal circuit including medial intraparietal area and dorsal premotor area (Filimon, 2010; Di Bono et al., 2015). Notably, the MLN is linked to a parietal network: Consistent with the ATOM theory (Walsh, 2003; Buetti and Walsh, 2009), the brain regions dedicated to number processing and to reach-to-grasp movements are closely linked by a generalized magnitude system, which transforms quantitative information into actions. In this connection, it would be important to consider the neural mechanism linking number and reaching movement. Functional neuroimaging studies will help to clarify the differential contribution of the reaching and the grasping components to number processing in action execution.

Embodied Number Processing

Our evidence, highlighting spatial and temporal properties of finger movements responding to numbers, fits with the embodied theory of numerical representation. From this perspective, numbers are not abstract, but embodied, i.e., rooted in bodily experiences. The way in which we use our bodies to act can influence our cognition (for an overview see Wilson, 2002; Barsalou, 2008; Patro et al., 2015). For example, the sensory-motor activations which occur during learning shape the newly learnt representation (Fischer and Zwaan, 2008; Fischer, 2012). Since number acquisition usually implies concomitant body-movements, like finger counting (Brissaud, 1992; Domahs et al., 2010), such embodied space-motor-number relations have also been used for training. For example it has been shown that playing games eliciting an embodied experience of the spatial layout of the MNL improves numerical competences (Ramani and Siegler, 2008; Whyte and Bull, 2008; Siegler and Ramani, 2009).

The challenging perspective of embodied cognition offers a stimulating approach to the study of mathematical competences. The analysis of finger movements indeed is considered a powerful method to assess numerical representations (Dotan and Dehaene, 2013). Paradigms focused on finger trajectories could be used to assess mental computations, and might offer a diagnostic instrument for measuring both normal and pathological development of mathematical competences (Booth and Siegler, 2006).

CONCLUSION

This study aimed at deepens our knowledge on the link between spatial numerical association and action execution. From an evolutive perspective, it could extend existing evidence on the origin of the spatial numerical association (de Hevia and Spelke, 2009; de Hevia et al., 2012; McCrink and Opfer, 2014; Nuerk et al., 2015; Rugani and de Hevia, 2016; Möhrling et al., 2017). Moreover, it could allow to clarify whether and how symbolic and non-symbolic numbers (see Rugani et al., 2017 for a definition) affect the sensorimotor transformations related to the motor control of the hand. This is particularly relevant considering the role of finger counting in number processing (Di Luca et al., 2006; Fischer, 2008; Sixtus et al., 2017), which survives in adults and seems to help the associations between numbers and hand actions (Hatano et al., 1977; Hubbard et al., 2005). The important relation between finger counting and mathematical abilities is scientifically documented. Abacus experts spontaneously move their hands while solving arithmetic calculation (Hatano et al., 1977). The manumerical cognition hypothesis (Hubbard et al., 2005) claims that this relation could explicate why dyscalculia, left-right confusion and finger agnosia often co-occur in the Gerstmann syndrome.

Last but not least, our protocol – based on a self-motivating and engaging task – would allow to investigate numerical cognition and its relation with space without the anxiety which usually disadvantages children with problems in mathematical comprehension (for the use of an innovative approach on

mathematical learning with Touchscreen Tablets, see Duijzer et al., 2017). Highly math-anxious persons have a common and strong tendency to avoid math, which reduces their possibility to increase their math competences (Ashcraft, 2002).

This paradigm, characterized by a new and not-overlearned task, could therefore be used to study the relation between number processing and motor action over development, before and during mathematical learning. Movements kinematics not only provides an accurate measure of the association between numbers and actions, but could also offer a novel tool for the diagnosis of potential mathematical deficits.

AUTHOR CONTRIBUTIONS

RR and LS designed the study, created the stimuli, collected, analyzed the data, and wrote the manuscript. LS did the statistical

analyses. RR, SB, and LS interpreted the data discussed the results. SB created the figures.

FUNDING

This work was supported by SIR grant (Scientific Independence of Young Researchers – No. RBSI141QKX) to LS and by “German Academic Exchange Service or DAAD (German: Deutscher Akademischer Austauschdienst),” Funding program: Research Stays for University Academics and Scientists, 2017 – No. 91644645 to RR.

ACKNOWLEDGMENTS

The authors would like to thank Elena Carlotta Borile for her help in conducting the experiments.

REFERENCES

- Andres, M., Davare, M., Pesenti, M., Olivier, E., and Seron, X. (2004). Number magnitude and grip aperture interaction. *Neuroreport* 15, 2773–2777.
- Andres, M., Olivier, E., and Badets, A. (2008a). Actions, words, and numbers: a motor contribution to semantic processing? *Curr. Dir. Psychol. Sci.* 17, 313–317. doi: 10.1111/j.1467-8721.2008.00597.x
- Andres, M., Ostry, D. J., Nicol, F., and Paus, T. (2008b). Time course of number magnitude interference during grasping. *Cortex* 44, 414–419. doi: 10.1016/j.cortex.2007.08.007
- Ansuini, C., Cavallo, A., Campus, C., Quarona, D., Koul, A., and Becchio, C. (2016). Are we real when we fake? Attunement to object weight in natural and pantomimed grasping movements. *Front. Hum. Neurosci.* 10:471. doi: 10.3389/fnhum.2016.00471
- Armbrüster, C., and Spijkers, W. (2006). Movement planning in prehension: do intended actions influence the initial reach and grasp movement? *Motor Control* 10, 311–329. doi: 10.1123/mcj.10.4.311
- Ashcraft, M. H. (2002). Math anxiety: personal, educational, and cognitive consequences. *Curr. Dir. Psychol. Sci.* 11, 181–185. doi: 10.1111/1467-8721.00196
- Badets, A., Andres, M., Di Luca, S., and Pesenti, M. (2007). Number magnitude potentiates action judgements. *Exp. Brain Res.* 180, 525–534. doi: 10.1007/s00221-007-0870-y
- Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645. doi: 10.1146/annurev.psych.59.103006.093639
- Betti, S., Castiello, U., and Sartori, L. (2015). Kick with the finger: symbolic actions shape motor cortex excitability. *Eur. J. Neurosci.* 42, 2860–2866. doi: 10.1111/ejn.13067
- Booth, J. L., and Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Dev. Psychol.* 42, 189–201. doi: 10.1037/0012-1649.41.6.189
- Brissiaud, R. (1992). “A tool for number construction: finger symbol sets,” in *Pathways to Number: Children’s Developing Numerical Abilities*, eds J. Bideaud, C. Meljac, and J.-P. Fischer (Hillsdale, NJ: Erlbaum), 41–65.
- Brouwer, A. M., Georgiou, I., Glover, S., and Castiello, U. (2006). Adjusting reach to lift movements to sudden visible changes in target’s weight. *Exp. Brain Res.* 173, 629–636. doi: 10.1007/s00221-006-0406-x
- Bueti, D., and Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1831–1840. doi: 10.1098/rstb.2009.0028
- Castiello, U. (1996). Grasping a fruit: selection for action. *J. Exp. Psychol. Hum. Percept. Perform.* 22, 582–603. doi: 10.1037/0096-1523.22.3.582
- Castiello, U. (2005). The neuroscience of grasping. *Nat. Rev. Neurosci.* 6, 726–736. doi: 10.1038/nrn1744
- Cheng, X., Ge, H., Andoni, D., Ding, X., and Fan, Z. (2015). Composite body movements modulate numerical cognition: evidence from the motion numerical compatibility effect. *Front. Psychol.* 6:1692. doi: 10.3389/fpsyg.2015.01692
- Chieffi, S., and Gentilucci, M. (1993). Coordination between the transport and the grasp components during prehension movements. *Exp. Brain Res.* 94, 471–477. doi: 10.1007/BF00230205
- Chiou, R. Y. C., Chang, E. C., Tzeng, O. J. L., and Wu, D. H. (2009). The common magnitude code underlying numerical and size processing for action but not for perception. *Exp. Brain Res.* 194, 553–562. doi: 10.1007/s00221-009-1730-8
- Chiou, R. Y. C., Wua, D. H., Tzeng, O. J. L., Hung, D. L., and Chang, E. C. (2012). Relative size of numerical magnitude induces a size-contrast effect on the grip scaling of reach-to-grasp movements. *Cortex* 48, 1043–1051. doi: 10.1016/j.cortex.2011.08.001
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edn. Hillsdale, NJ: Lawrence Erlbaum Associates.
- D’Amico, M., and Ferrigno, G. (1990). Technique for the evaluation of derivatives from noisy biomechanical displacement data using a model-based bandwidth-selection procedure. *Med. Biol. Eng. Comput.* 28, 407–415. doi: 10.1007/BF02441963
- D’Amico, M., and Ferrigno, G. (1992). Comparison between the more recent techniques for smoothing and derivative assessment in biomechanics. *Med. Biol. Eng. Comput.* 30, 193–204. doi: 10.1007/BF02446130
- de Hevia, M. D., and Spelke, E. S. (2009). Spontaneous mapping of number and space in adults and young children. *Cognition* 110, 198–207. doi: 10.1016/j.cognition.2008.11.003
- de Hevia, M. D., Vanderslice, M., and Spelke, E. S. (2012). Cross dimensional mapping of number, length and brightness by preschool children. *PLoS One* 7:e35530. doi: 10.1371/journal.pone.0035530
- Dehaene, S. (2011). *The Number Sense: How the Mind Creates Mathematics, Revised and Updated*. New York, NY: Oxford University Press.
- Dehaene, S., Bossini, S., and Giraux, P. (1993). The mental representation of parity and number magnitude. *J. Exp. Psychol. Gen.* 122, 371–396. doi: 10.1037/0096-3445.122.3.371
- Di Bono, M. G., Begliomini, C., Castiello, U., and Zorzi, M. (2015). Probing the reaching-grasping network in humans through multivoxel pattern decoding. *Brain Behav.* 5:e00412. doi: 10.1002/brb3.412
- Di Luca, S., Granà, A., Semenza, C., Seron, X., and Pesenti, M. (2006). Finger-digit compatibility in Arabic numeral processing. *Q. J. Exp. Psychol.* 59, 1648–1663. doi: 10.1080/17470210500256839
- Domahs, F., Moeller, K., Huber, S., Willmes, K., and Nuerk, H. C. (2010). Embodied numerosity: implicit hand-based representations influence symbolic number

- processing across cultures. *Cognition* 116, 251–266. doi: 10.1016/j.cognition.2010.05.007
- Dotan, D., and Dehaene, S. (2013). How do we convert a number into a finger trajectory? *Cognition* 129, 512–529. doi: 10.1016/j.cognition.2013.07.007
- Duijzer, C. A., Shayan, S., Bakker, A., Van der Schaaf, M. F., and Abrahamson, D. (2017). Touchscreen tablets: coordinating action and perception for mathematical cognition. *Front. Psychol.* 8:144. doi: 10.3389/fpsyg.2017.00144
- Eastough, D., and Edwards, M. G. (2007). Movement kinematics in prehension are affected by grasping objects of different mass. *Exp. Brain Res.* 176, 193–198. doi: 10.1007/s00221-006-0749-3
- Faulkenberry, T. J. (2014). Hand movements reflect competitive processing in numerical cognition. *Can. J. Exp. Psychol.* 68, 147–151. doi: 10.1037/cep0000021
- Faulkenberry, T. J., Cruise, A., Lavro, D., and Shaki, S. (2016). Response trajectories capture the continuous dynamics of the size congruity effect. *Acta Psychol.* 163, 114–123. doi: 10.1016/j.actpsy.2015.11.010
- Faulkenberry, T. J., Cruise, A., and Shaki, S. (2017). Reversing the manual digit bias in two-digit number comparison. *Exp. Psychol.* 64, 191–204. doi: 10.1027/1618-3169/a000365
- Fias, W., and Fischer, M. H. (2005). “Spatial representation of number,” in *Handbook of Mathematical Cognition*, ed. J. Campbell (New York, NY: Psychology Press), 43–54.
- Filimon, F. (2010). Human cortical control of hand movements: parietofrontal networks for reaching, grasping, and pointing. *Neuroscientist* 16, 388–407. doi: 10.1177/1073858410375468
- Fischer, M. H. (2001). Number processing induces spatial performance biases. *Neurology* 57, 822–826. doi: 10.1212/WNL.57.5.822
- Fischer, M. H. (2003). Spatial representations in number processing—Evidence from a pointing task. *Vis. Cogn.* 10, 493–508. doi: 10.1080/13506280244000186
- Fischer, M. H. (2008). Finger counting habits modulate spatial-numerical associations. *Cortex* 44, 386–392. doi: 10.1016/j.cortex.2007.08.004
- Fischer, M. H. (2012). A hierarchical view of grounded, embodied, and situated numerical cognition. *Cogn. Process.* 13, 161–164. doi: 10.1007/s10339-012-0477-5
- Fischer, M. H., and Zwaan, R. A. (2008). Embodied language: a review of the role of the motor system in language comprehension. *Q. J. Exp. Psychol.* 61, 825–850. doi: 10.1080/17470210701623605
- Gallese, V., and Lakoff, G. (2005). The brain's concepts: the role of the sensory-motor system in conceptual knowledge. *Cogn. Neuropsychol.* 22, 455–479. doi: 10.1080/02643290442000310
- Galton, F. (1880). Visualised numerals. *Nature* 21, 252–256. doi: 10.1038/021252a0
- Gentilucci, M., Benuzzi, F., Bertolani, L., Daprati, E., and Gangitano, M. (2000). Recognising a hand by grasp. *Cogn. Brain Res.* 9, 125–135. doi: 10.1016/S0926-6410(99)00049-X
- Gentilucci, M., Campione, G. C., De Stefani, E., and Innocenti, A. (2012). Is the coupled control of hand and mouth postures precursor of reciprocal relations between gestures and words? *Behav. Brain Res.* 233, 130–140. doi: 10.1016/j.bbr.2012.04.036
- Gevers, W., Reynvoet, B., and Fias, W. (2003). The mental representation of ordinal sequences is spatially organized. *Cognition* 87, B87–B95. doi: 10.1016/S0010-0277(02)00234-2
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Hillsdale, NJ: Lawrence Erlbaum.
- Girelli, L., Perrone, G., Croccolo, F., Roman, E. H., Bricolo, E., Mancin, M., et al. (2016). Manual actions cover symbolic distances at different speed. *Acta Psychol.* 169, 56–60. doi: 10.1016/j.actpsy.2016.05.002
- Grade, S., Lefèvre, N., and Pesenti, M. (2013). Influence of gaze observation on random number generation. *Exp. Psychol.* 60, 122–130. doi: 10.1027/1618-3169/a000178
- Hartmann, M., Grabherr, L., and Mast, F. W. (2012). Moving along the number line: interactions between whole-body motion and numerical cognition. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 1416–1427. doi: 10.1037/a0026706
- Hatano, G., Miyake, Y., and Binks, M. G. (1977). Performance of expert abacus operators. *Cognition* 5, 47–55. doi: 10.1016/0010-0277(77)90016-6
- Hubbard, E. M., Piazza, M., Pinel, P., and Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nat. Rev. Neurosci.* 6, 435–448. doi: 10.1038/nrn1684
- Ishihara, M., Jacquin-Courtois, S., Flory, V., Salemm, R., Imanaka, K., and Rossetti, Y. (2006). Interaction between space and number representations during motor preparation in manual aiming. *Neuropsychologia* 44, 1009–1016. doi: 10.1016/j.neuropsychologia.2005.11.008
- Jakobson, L. S., and Goodale, M. A. (1991). Factors affecting higher-order movement planning: a kinematic analysis of human prehension. *Exp. Brain Res.* 86, 199–208. doi: 10.1007/BF00231054
- Jeannerod, M. (1981). “Intersegmental coordination during reaching at natural visual objects,” in *Attention and Performance IX*, eds J. Long and A. Baddeley (Hillsdale, NY: Erlbaum), 153–168.
- Krishnan-Barman, S., Forbes, P. A., and Hamilton, A. F. C. (2017). How can the study of action kinematics inform our understanding of human social interaction? *Neuropsychologia* 105, 101–110. doi: 10.1016/j.neuropsychologia.2017.01.018
- Lindemann, O., Abolafia, J. M., Girardi, G., and Bekkering, H. (2007). Getting a grip on numbers: numerical magnitude priming in object grasping. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 1400–1409. doi: 10.1037/0096-1523.33.6.1400
- Loetscher, T., Schwarz, U., Schubiger, M., and Brugger, P. (2008). Head turns bias the brain's internal random generator. *Curr. Biol.* 18, R60–R62. doi: 10.1016/j.cub.2007.11.015
- McCrink, K., and Opfer, J. E. (2014). Development of spatial-numerical associations. *Curr. Dir. Psychol. Sci.* 23, 439–445. doi: 10.1177/0963721414549751
- Möhring, W., Ishihara, M., Curiger, J., and Frick, A. (2017). Spatial-numerical associations in first-graders: evidence from a manual-pointing task. *Psychol. Res.* doi: 10.1007/s00426-017-0904-4 [Epub ahead of print].
- Moretto, G., and di Pellegrino, G. (2008). Grasping numbers. *Exp. Brain Res.* 188, 505–515. doi: 10.1007/s00221-008-1386-9
- Namdar, G., Tzelgov, J., Algom, D., and Ganel, T. (2014). Grasping numbers: evidence for automatic influence of numerical magnitude on grip aperture. *Psychon. Bull. Rev.* 21, 830–835. doi: 10.3758/s13423-013-0550-9
- Nuerk, H.-C., Patro, K., Cress, U., Schild, U., Friedrich, C. K., and Goebel, S. M. (2015). How space-number associations may be created in preliterate children: six distinct mechanisms. *Front. Psychol.* 6:215. doi: 10.3389/fpsyg.2015.00215
- Patro, K., Nuerk, H. C., and Cress, U. (2015). Does your body count? Embodied influences on the preferred counting direction of preschoolers. *J. Cogn. Psychol.* 27, 413–425. doi: 10.1080/20445911.2015.1008005
- Ramani, G. B., and Siegler, R. S. (2008). Promoting broad and stable improvements in low-income children's numerical knowledge through playing number board games. *Child Dev.* 79, 375–394. doi: 10.1111/j.1467-8624.2007.01131.x
- Rugani, R., Betti, S., Ceccarini, F., and Sartori, L. (2017). Act on numbers: numerical magnitude influences selection and kinematics of finger movement. *Front. Psychol.* 8:1481. doi: 10.3389/fpsyg.2017.01481
- Rugani, R., and de Hevia, M. D. (2016). Number-space associations without language. Evidence from preverbal human infants and non-human animal species. *Psychon. Bull. Rev.* 24, 352–369. doi: 10.3758/s13423-016-1126-2
- Rugani, R., and Sartori, L. (2016). Numbers in action. *Front. Hum. Neurosci.* 10:388. doi: 10.3389/fnhum.2016.00388
- Santens, S., Goossens, S., and Verguts, T. (2011). Distance in motion: response trajectories reveal the dynamics of number comparison. *PLoS One* 6:e25429. doi: 10.1371/journal.pone.0025429
- Schuboe, A., Maldonado, A., Stork, S., and Beetz, M. (2008). “Subsequent actions influence motor control parameters of a current grasping action,” in *Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN, Yokohama*, 389–394. doi: 10.1109/ROMAN.2008.4600697
- Schwarz, W., and Keus, I. M. (2004). Moving the eyes along the mental number line: comparing SNARC effects with saccadic and manual responses. *Percept. Psychophys.* 66, 651–664. doi: 10.3758/BF03194909
- Shaki, S., and Fischer, M. H. (2014). Random walks on the mental number line. *Exp. Brain Res.* 232, 43–49. doi: 10.1007/s00221-013-3718-7
- Siegler, R. S., and Ramani, G. B. (2009). Playing linear number board games—but not circular ones—improves low-income preschoolers' numerical understanding. *J. Educ. Psychol.* 101, 545–560. doi: 10.1037/a0014239
- Sixtus, E., Fischer, M. H., and Lindemann, O. (2017). Finger posing primes number comprehension. *Cogn. Process.* 18, 237–248. doi: 10.1007/s10339-017-0804-y
- Smeets, J. B. J., and Brenner, E. (1999). A new view on grasping. *Motor Control* 3, 237–271. doi: 10.1123/mcj.3.3.237

- Song, J.-H., and Nakayama, K. (2008). Numeric comparison in a visually-guided manual reaching task. *Cognition* 106, 994–1003. doi: 10.1016/j.cognition.2007.03.014
- Tucker, M., and Ellis, R. (1998). On the relations between seen objects and components of potential actions. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 830–846. doi: 10.1037/0096-1523.24.3.830
- Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends Cogn. Sci.* 7, 483–488. doi: 10.1016/j.tics.2003.09.002
- Weir, P. L., MacKenzie, C. L., Marteniuk, R. G., Cargoe, S. L., and Frazer, M. B. (1991). The effects of object weight on the kinematics of prehension. *J. Mot. Behav.* 23, 192–204. doi: 10.1080/00222895.1991.10118362
- Whyte, J. C., and Bull, R. (2008). Number games, magnitude representation, and basic number skills in preschoolers. *Dev. Psychol.* 44, 588–596. doi: 10.1037/0012-1649.44.2.588
- Wilson, M. (2002). Six views of embodied cognition. *Psychon. Bull. Rev.* 9, 625–636. doi: 10.3758/BF03196322
- Winter, B., Matlock, T., Shaki, S., and Fischer, M. H. (2015). Mental number space in three dimensions. *Neurosci. Biobehav. Rev.* 57, 209–219. doi: 10.1016/j.neubiorev.2015.09.005
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Rugani, Betti and Sartori. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



How Deep Is Your SNARC? Interactions Between Numerical Magnitude, Response Hands, and Reachability in Peripersonal Space

Johannes Lohmann^{1*}, Philipp A. Schroeder^{2,3}, Hans-Christoph Nuerk^{3,4,5},
Christian Plewnia^{2,6} and Martin V. Butz^{1,6}

¹ Cognitive Modeling, Department of Computer Science, University of Tübingen, Tübingen, Germany, ² Department of Psychiatry and Psychotherapy, Neurophysiology and Interventional Neuropsychiatry, University of Tübingen, Tübingen, Germany, ³ Department of Psychology, University of Tübingen, Tübingen, Germany, ⁴ Leibniz-Institut für Wissensmedien (IWM), Tübingen, Germany, ⁵ LEAD Graduate School, University of Tübingen, Tübingen, Germany, ⁶ Werner Reichardt Centre for Integrative Neuroscience, Tübingen, Germany

OPEN ACCESS

Edited by:

Ann Dowker,
University of Oxford, United Kingdom

Reviewed by:

Martin H. Fischer,
University of Potsdam, Germany
Annemie Desoete,
Ghent University, Belgium

*Correspondence:

Johannes Lohmann
johannes.lohmann@uni-tuebingen.de

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 21 December 2017

Accepted: 12 April 2018

Published: 01 May 2018

Citation:

Lohmann J, Schroeder PA,
Nuerk H-C, Plewnia C and Butz MV
(2018) How Deep Is Your SNARC?
Interactions Between Numerical
Magnitude, Response Hands,
and Reachability in Peripersonal
Space. *Front. Psychol.* 9:622.
doi: 10.3389/fpsyg.2018.00622

Spatial, physical, and semantic magnitude dimensions can influence action decisions in human cognitive processing and interact with each other. For example, in the spatial-numerical associations of response code (SNARC) effect, semantic numerical magnitude facilitates left-hand or right-hand responding dependent on the small or large magnitude of number symbols. SNARC-like interactions of numerical magnitudes with the radial spatial dimension (depth) were postulated from early on. Usually, the SNARC effect in any direction is investigated using fronto-parallel computer monitors for presentation of stimuli. In such 2D setups, however, the metaphorical and literal interpretation of the radial depth axis with seemingly close/far stimuli or responses are not distinct. Hence, it is difficult to draw clear conclusions with respect to the contribution of different spatial mappings to the SNARC effect. In order to disentangle the different mappings in a natural way, we studied parametrical interactions between semantic numerical magnitude, horizontal directional responses, and perceptual distance by means of stereoscopic depth in an immersive virtual reality (VR). Two VR experiments show horizontal SNARC effects across all spatial displacements in traditional latency measures and kinematic response parameters. No indications of a SNARC effect along the depth axis, as it would be predicted by a direct mapping account, were observed, but the results show a non-linear relationship between horizontal SNARC slopes and physical distance. Steepest SNARC slopes were observed for digits presented close to the hands. We conclude that spatial-numerical processing is susceptible to effector-based processes but relatively resilient to task-irrelevant variations of radial-spatial magnitudes.

Keywords: SNARC effect, theory of magnitude, embodied numerical cognition, virtual reality, motion capture

INTRODUCTION

Relational inference is fundamental for adaptive behavior control. Catching a flying object requires an estimate of the hand position in space and time as well as the velocity of the object. Even simple grasping movements require a thorough estimate of the distance between the target object and the own body. From a conceptual point of view, these estimates are similar since they all require

magnitude judgments – in time, space, and the respective derivatives thereof, that is, speed and acceleration. There is indeed evidence for a common metric involved in the representation of time, space, and quantity. According to a *theory of magnitude* (ATOM; Walsh, 2003), this metric evolves from the sensorimotor system and resides primarily in the parietal cortices of the brain (Walsh, 2003; Buetti and Walsh, 2009; Cohen-Kadosh and Dowker, 2015). According to ATOM, the common magnitude metric emerges in the service of action control and develops into a general magnitude system that can be used to represent arbitrary quantities, for instance, in terms of numbers. Hence, ATOM can account for the apparent overlap of magnitude processing across modalities.

One example for such an overlap with respect to space and magnitude is the well-studied spatial-numerical associations of response codes (SNARC) effect. The SNARC effect shows a strong interaction between directional spatial information in the left-hand/right-hand of responding and the numerical magnitude information as semantically displayed in numerical symbols (Dehaene et al., 1993; Wood et al., 2008). In simple judgment tasks on numerical magnitude or parity, small numbers are faster responded to with the left hand than with the right hand, and vice versa for large numbers. SNARC effects can be obtained with different response systems such as hands, eyes, or feet (Fischer et al., 2003; Schwarz and Müller, 2006; Hesse and Bremmer, 2016), for different modalities and number notations (Nuerk et al., 2005), and SNARC effects can also influence overt action decisions, which nicely demonstrates the relevance of the metrical overlap for action coordination in more or less naturalistic settings (Shaki and Fischer, 2014; Schroeder and Pfister, 2015). Furthermore, interactions have been documented between the spatial information triggered by different magnitudes such as auditory and visual intensity (Fairhurst and Deroy, 2017) or by number and musical pitch in both factorial designs (Weis et al., 2016) as well as in dual-task situations (Fischer et al., 2013). These findings are consistent with the assumption of a common magnitude representation, which is assumed to be located within the horizontal segment of the intraparietal sulcus (Dehaene et al., 2003), and highlight the overlap between numerical and spatial cognition.

Usually, the SNARC effect has been interpreted in terms of a mental number line, oriented horizontally, with small numbers represented to the left of large numbers. However, different studies have shown that spatial associations of numerical magnitude are not restricted to the horizontal dimension, but can be extended to the vertical and radial dimension as well. ATOM can account for the existence of all of these mappings, by stating that magnitudes are flexibly mapped on spatial dimensions involved in the task. According to this line of reasoning, for instance radial SNARC effects can arise because nearby space corresponds with a small movement amplitude and thus shares the meaning of small magnitudes with small numbers. This implies that spatial-numerical mappings are more flexible and are not restricted to a single, horizontal representation. However, ATOM does not provide a prediction, which kind of mapping is applied under which circumstances. From an anticipatory behavior

control perspective (e.g., Hoffmann, 1993, 2003), the application of a certain mapping should not occur automatically, but should be driven by task relevance. Accordingly, we pursued two broad aims with the current study. First, we wanted to corroborate further evidence for a sensorimotor grounding of SNARC effects. Second, we wanted to investigate the situatedness of spatial-numerical mappings in task-relevant and task-irrelevant spatial dimensions. In order to do so, an experimental setup would be desirable that allows to contrast different spatial axes within the same environment, and which provides a natural user interface. Hence, we realized a SNARC setup in an immersive virtual reality (VR), combined with online motion capture.

How Deep Is the SNARC Effect?

Already the first scientific description of SNARC-like effects included rather diverse (and partially complicated) introspective self-reports of mental number lines wandering through space, also extending to the radial depth dimension (Galton, 1880). However, to date, only relatively few studies have tested other spatial directions than the horizontal left-to-right plane, or even tested combinatory-factorial experimental designs to investigate interactions between the potentially available horizontal, vertical, or radial (distance-based or sagittal) SNARC effects (for an exhaustive review, see Winter et al., 2015). When studied in isolation, spatial-numerical associations were observed (at least in Western cultures and besides the left-to-right direction) for lower-hand vs. upper-hand (but not feet) responses from bottom-to-top (Hartmann et al., 2012; Wiemers et al., 2017) and also when responses were mapped from back-to-front (i.e., vertical in the sense of close/far from the body; Ito and Hatta, 2004; Shaki and Fischer, 2012). However, in traditional setups using fronto-parallel two-dimensional computer monitors for presentation of stimuli, the metaphorical and literal interpretation of close/far (along with the linguistic declaration thereof) are not necessarily distinct. This is problematic because also vertical labels and horizontal response arrangements can produce spatial-numerical associations (Holmes and Lourenco, 2011). Since spatial associations in different spatial dimensions could have different cognitive origins (Winter et al., 2015; Wiemers et al., 2017), it is not clear whether which dimensions would produce an effect or how the different spatial and numerical magnitudes would interact. Nevertheless, at least semantically, there seems to be an association between close-small and far-large (Santens and Gevers, 2008). Results implying the presence of radial SNARC effects circulating the body have been reported by Marghetis and Youngstrom (2014). In their study, participants had to judge the magnitude of single-digit numbers by stepping forward or backward. Marghetis and Youngstrom (2014) compared performance in the magnitude judgment for whole numbers (1 to 9, except 5) and integers (−9 to 9, without 0). In the latter task, a SNARC-like pattern emerged, with backward responses being faster for negative numbers, and forward responses being faster for positive numbers. However, if the stimulus-set only contained positive numbers, no association between magnitude and movement direction was observed.

As it was pointed out by Winter et al. (2015), the multitude of flexible spatial-numerical mappings and their task-dependency renders numerical cognition highly situated. Furthermore, the reviewed findings imply that numerical cognition builds upon a rich spatial representation, which can also exceed implicit directional SNARC effects on other, explicit linkages and in effects of spatial extension (Patro et al., 2014; Cipora et al., 2015). According to ATOM, this spatial representation is the same that is used for behavior control. Indeed, there is some evidence for a close relation between the multisensory spatial mappings used to represent the space surrounding the body – the so-called peripersonal space (e.g., Holmes and Spence, 2004) – and numerical space. Longo and Lourenco (2010) investigated whether biases of lateralized attention within peripersonal space also apply to numerical cognition. In pen-and-paper line bisection, a small leftward bias is typically observed for lines close to the body, which reverse to a rightward bias with increasing distance. Longo and Lourenco (2010) observed the same bias and a similar effect of physical distance if participants had to bisect number pairs. Furthermore, the size of both biases was highly correlated on an individual level. This implies a close coupling of the representation of physical and numerical space. Further evidence for this coupling was provided by Patro et al. (2015), who showed that counting directions in preliterate children are emphasized in peripersonal space. Moreover, it could be that the flexible change between different egocentric and allocentric perspectives and the transformations between peripersonal and extrapersonal spaces contribute to the effects of embodied numerical learning paradigms (Dackermann et al., 2017). Together, these findings imply a highly flexible representation, which is used to map numbers and space, and that this representation is closely tied to the representation of physical space, which is grounded in sensorimotor experience.

Further evidence for the sensorimotor grounding of numerical representations proposed by ATOM comes from studies implying SNARC-like number-action links. For instance, it has been shown that numerical magnitude can afford compatible grip apertures (Andres et al., 2004). In this study, participants had to close or open their hand in response to a digit's parity. Closure was faster in case of small digits, while opening was faster in case of large digits. In a similar vein, it has been shown that large digits afford power grasps, while small digits afford precision grasps – even if the numerical magnitude is not task relevant (Lindemann et al., 2007).

Regarding *interactions* between SNARC effects with different spatial codes in the response dimension, only few studies have previously pitted different spatial dimensions against each other, and if they did so, diagonal response mappings were used (Gevers et al., 2006; Holmes and Lourenco, 2011, 2012). Noteworthy, the perceptual presentation of semantic magnitudes (in form of Arabic single-digits) was mostly carried out using two-dimensional stimuli on flat computer displays, varying only the spatial response dimension in horizontal, vertical, or radial direction. However, regarding SNARC effects with different spatial codes in terms of visual-perceptual presentation, to the best of our knowledge, there was no systematic investigation up to now. In the present study, we investigated whether the possibility

of concurrent extensions on the two-dimensional fronto-parallel and three-dimensional proximal-distal plane yields a more complicated and possibly interacting scheme of a single-digit's spatial associations.

From the available literature, two main hypotheses can be formulated. If associations between spatial and numerical magnitudes are driven by direct mappings of perceptual magnitudes on spatial directions, there should be crossmodal interactions at the level of the theoretical core magnitude system, as it was also repeatedly found for other magnitude dimensions (Fischer et al., 2013; Weis et al., 2016; Fairhurst and Deroy, 2017) or for the direct comparison between semantic magnitude and physical extension in the size congruity effect (Henik and Tzelgov, 1982). The interactions should be detectable even if different psychophysical scales for the distinct spatial dimensions might result in different magnitude weights (see Winter et al., 2015 for a similar argument). However, considering the previous results implying a relation between physical and numerical space (Longo and Lourenco, 2010; Patro et al., 2015; Dackermann et al., 2017), changes in reachability or the transition from peri- to extra-personal space might result in a more complex modulation of SNARC effects along the radial axis.

Embedding Numerical Cognition in Virtual Reality: The Present Study

In the present study, we introduce a VR scenario to systematically investigate the interaction between perceptual distance and horizontal SNARC effects. Compared to classic, fronto-parallel display setups used to study SNARC effects, VR allows to vary perceptual distance without confusion with the vertical dimension in a three-dimensional stereoscopic simulation. This allows the combination of a horizontal response mapping with stimulus presentation on the radial axis and hence, spatial codes in the response and presentation dimensions can be varied experimentally. Furthermore, the incorporation of online motion capture allows the implementation of a natural, continuous response mode, as well as sensorimotor exploration of the task space.

Two distinct procedures were carried out in the present research. First, although the simulation in VR already includes stereoscopic 3D images (using the Oculus Rift® DK2 head-mounted display), we furthermore included a sensorimotor exploration phase prior to the actual SNARC experiment to provide an immediate experience of peripersonal space in the VR test environment, and possibly adjust for individual differences in overestimation of perceived reachability (Fischer, 2005). To that end, in our implementation, the Leap Motion® near-infrared sensor was used to track and stream hand movements to the VR scenario. Such setups, which allow participants to explore the VR with a body representation, have previously shown to increase the degree of immersion and spatial perception within the VR (Mohler et al., 2008; Linkenauger et al., 2015). Furthermore, there is evidence that the distinction between peri- and extra-personal space remains valid in suitable VR setups (Gamberini et al., 2008). Second, in order to obtain an action-related, kinematic measure of the response activation during the task, and closely following the results of

contralateral motor activation in the incongruent conditions of the SNARC effect (e.g., Keus et al., 2005), we used a slightly different effector response than in previous studies, which had mostly utilized response box key presses. More precisely, the response mode in the current experiment was realized by asking participants to close their hands, which were positioned at a fixed and comfortable distance in the VR display. Thus, this response modality further allowed for continuous response activation in conflicting conditions, next to the established assessment of SNARC effects by means of response times (RTs) and regression coefficient analysis. In a comparable VR setup using the same equipment, we were previously able to reproduce the behavioral bias for food stimuli (Schroeder et al., 2016).

To conclude the motivation for the current study, the concurrent assessment of SNARC effects in the three-dimensional VR environment – including spatial displacements within and outside reachable space – allowed us to investigate interactions between spatial-numerical mappings on radial and horizontal axes. A direct mapping approach would predict a linear relationship between numerical magnitude and spatial magnitude on the radial axis. Precisely, in this case, left-side responding should be faster for small semantic digits (horizontal SNARC), but also for digits appearing closer to the participants (radial SNARC), and vice versa for right-hand responding. If SNARC effects are tied to spatial representations used in behavior control, as proposed by ATOM, a non-linear relation between numerical magnitude and radial distance – indicating effects of reachability or the transition from peri- to extra-personal space – seems more likely. In order to investigate these two hypotheses, we had participants perform a magnitude judgment with respect to digits appearing at different distances on the radial axis within or outside peripersonal space. In a first experiment, we analyzed the interactions between the SNARC effect and physical distance by applying 10 equidistant spatial displacements. In a second experiment, we focused on the four most relevant displacements identified in the first study.

MATERIALS AND METHODS

Participants

Sixteen students from the University of Tübingen participated in the first experiment (seven females). Their age ranged from 19 to 30 years ($M = 22.3$, $SD = 2.8$). All participants were right-handed and had normal or corrected-to-normal vision. Participants provided informed consent and received either course credit or a monetary compensation for their participation. For the second experiment, another 16 participants were recruited (10 females), none of whom participated in the first experiment. Their age ranged from 19 to 29 years ($M = 22.0$, $SD = 2.8$). Again, all participants were right-handed and had normal or corrected-to-normal vision. They provided informed consent and were compensated with course credit or money for their participation. Both experiments were conducted in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

Apparatus

To immerse participants in the VR, they were equipped with an Oculus Rift® DK2 stereoscopic head mounted display (HMD; Oculus VR LLC, Menlo Park, CA, United States). Motion tracking of hand movements was realized with a LeapMotion® near-infrared sensor (LeapMotion, Inc., San Francisco, CA, United States; SDK version 3.1.3). The LeapMotion® sensor provides positional information regarding the palm, wrist, and phalanges. This data can be used to render a hand model in VR. Furthermore, the API provides a measure of the hand closure of the respective hand, ranging between zero (open hand) and one (clenched fist). This measure of hand closure was used to determine the response in the SNARC task. A response was collected if the respective value was larger than 0.75. The whole experiment was implemented within the Unity® engine 5.5.0 using the C# interface provided by the API. To allow the experimenter to observe the scene and to assist the participants, the VR scene was rendered in parallel on the Oculus Rift and a computer screen.

Virtual Reality Setup

The VR setup put participants on a meadow surrounded by hills and various trees. A black plane covered with equally spaced white lines appeared in front of them. These lines indicated the displacements in the radial plane where the stimuli for the exploration and the magnitude judgment would appear. We chose these discrete distance indicators to make the distinction between reachable and not-reachable space even more salient. The distance between adjacent lines was about 10 cm. The center of the tracking range of the LeapMotion® sensor corresponded with the fourth line in the setup (second experiment: second line). The outer, radial limit of the tracking range was indicated by a cardboard box to provide participants with haptic feedback regarding the bounds of the task space (calibrated with the sixth/third visual horizontal line in the VR and with the tip of the middle finger with maximally extended arm). Please note that the interaction range was limited by the sensor range, which covered about 60 cm in depth and 50 cm from left to right, and not by the length of participants' arms. The real-world setup of the task space is shown in **Figure 1**. Instructions and feedback were presented on different text-fields, aligned at eye-height.

Procedure

At the beginning of the experiment, participants received a verbal instruction regarding the VR equipment. Then, the HMD was put on and the experiment started. In a first step, the scene was calibrated according to the participant's height and arm length, that is: the ground position was adjusted in a way that the hand appeared above the task space when it was stretched out to the outer bound of the reachable space (see **Figure 2**). In order to do so, participants had to stretch out their dominant right arm and place the tip of their middle finger on the top of a card box, which was placed at the border of the LeapMotion® sensor's tracking range. If necessary, the experimenter gently corrected the participant's seated position to assure that his or her arm

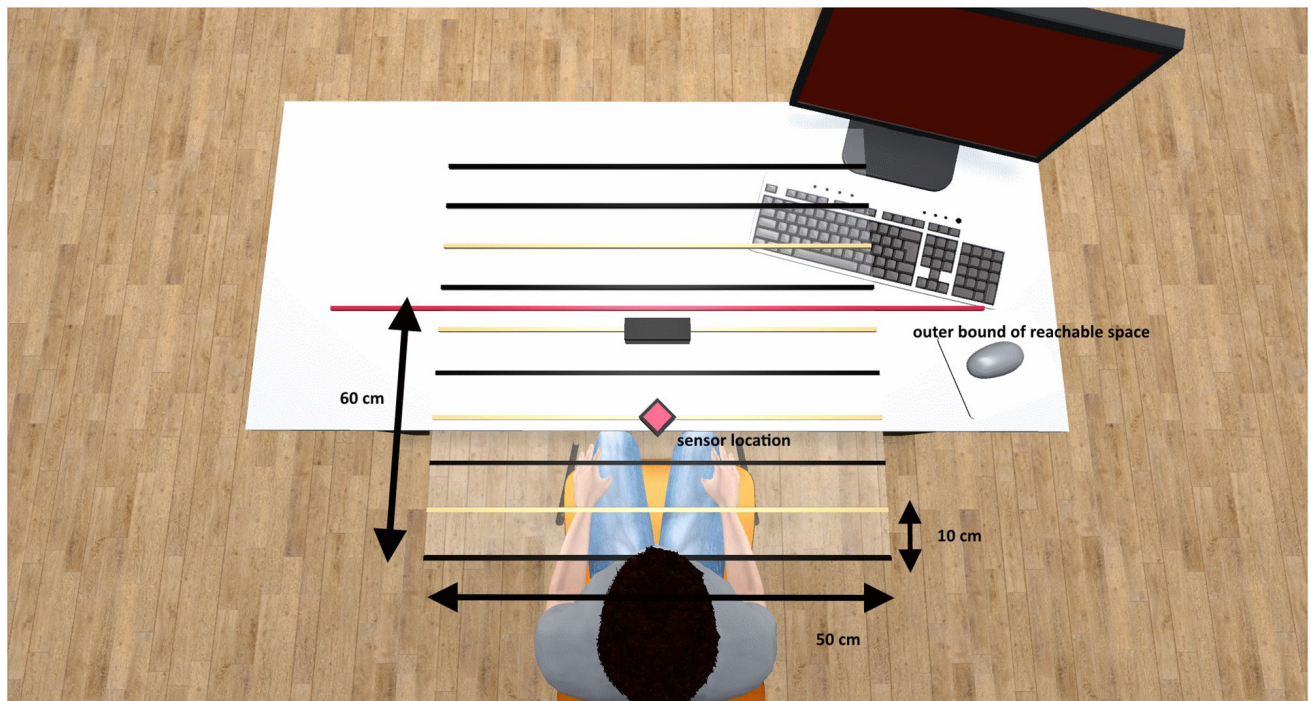


FIGURE 1 | Physical setup and extent of the task space. The red diamond indicates the sensor position and the red line indicates the outer bound of the reachable space. The lines correspond to the distance indicators in the VR environment. Yellow lines indicate the four distances that were applied for stimulus presentation in both experiments. Spacing between adjacent lines was about 10 cm.

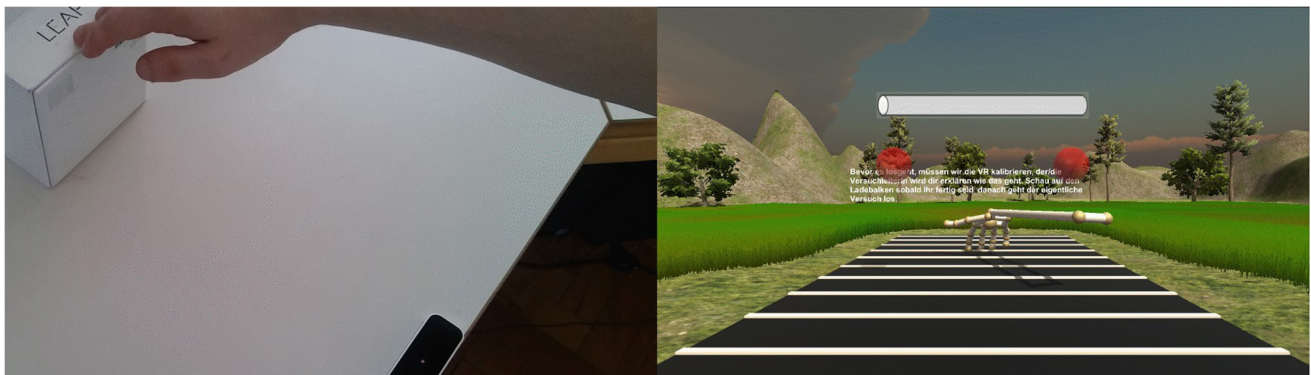


FIGURE 2 | The calibration setup. The virtual environment is shown in the right panel and the white lines correspond to the lines shown in **Figure 1**. At the beginning of the experiment, the height of the sensor origin was adjusted in such a way that a stretched arm (left panel) yielded the impression shown in the right panel. As it is shown in the right panel, four displacements were outside the reachable space. The red spheres indicate the initial hand positions for the magnitude judgment, in the calibration phase it was assured that these positions could be reached and held conveniently.

were maximally extended to reach the box (see **Figure 2**, left panel). Next, the experimenter adjusted the visual position of the virtual hand model to assure that the virtual hand appeared above the task space (see **Figure 2**, right panel). Furthermore, it was assured that the response hand positions for the SNARC task could be reached conveniently. This procedure ensured that participants experienced a standardized reachability limit in VR and the calibration furthermore reduced the influence of reaching range overestimations (Fischer, 2005).

The experiment consisted of two parts. First, participants performed an exploration task. This was intended to familiarize the participants with the sensorimotor mapping and to provide an experience of reachability. Second, participants performed two blocks of a magnitude judgment task within the VR. Participants could practice the magnitude judgment for 20 trials before the actual blocks started. Both tasks are described in detail below. After the experiment, participants were asked to complete a presence questionnaire (IPQ; Schubert et al., 2001). The whole

procedure took 90 to 120 min, including preparation and practice trials.

Sensorimotor Exploration Phase

To familiarize the participants with the sensorimotor mapping with respect to the different displacements and to enhance their depth perception, participants performed a reaching task within the VR. In this task (presented in the same environment as the later magnitude judgment task), colored spheres appeared at different spatial displacements, indicated by horizontal lines. Participants had to touch the spheres with the fingertips of their left or right hand. The color of the spheres indicated the requested hand: participants had to touch yellow spheres with their left hand and green spheres with their right hand. Upon touching, the spheres emitted a flashing burst. If participants touched the sphere with the correct hand, the flash was white. If they touched the sphere with the wrong hand, the flash was red.

If the spheres appeared at unreachable distances (displacements 7–10 in the first experiment, displacement 4 in the second experiment), participants were requested to press an accordingly labeled button (“too far,” German: “zu weit”) on the right side of the task space. The setup for the sensorimotor exploration task is shown in **Figure 3**. Participants had to perform 10 reaching movements per displacement, five with the left and five with the right hand, yielding 100 trials in the first experiment (10 different displacements) and 40 trials in the second experiment (four different displacements). The 10 repetitions per distance sampled the whole width of the task space, covering the left and right space. Participants had to perform ipsilateral as well as contralateral reaching movements. The order of presentation was randomized and error trials were not repeated. The performance in this task was not evaluated, because the exploration was only intended to familiarize the participants with the environment and to provide a behavioral experience of reachability and distance.

Magnitude Judgment Task

After the sensorimotor exploration phase, participants were requested to perform two blocks of a magnitude judgment task. Here, they had to repeatedly classify single-digits (1–4, 6–9) as being either smaller or larger than 5 by clenching their left or right fist. The response mapping varied between the two blocks: in one block, participants had to clench their right fist in case of digits larger than 5 and their left fist in case of digits smaller than 5. This mapping was reversed in the other block. The order of the response mapping was randomized.

Both blocks in both experiments consisted of 320 trials and each trial consisted of two parts. At the beginning of a trial, participants had to move their hands into initial positions, indicated by red, semi-transparent spheres, and located at the fourth (first experiment), or the second displacement (second experiment), respectively (see **Figures 2, 4**). If the palms were within the positions and the respective hands were open, the spheres turned green. Furthermore, participants had to center their field of view on a fixation cross located at the outer bound of the task space. The inner part of the fixation cross

turned green once the center of the visual field had been directed toward the fixation cross for at least 2000 ms (see **Figure 4**, left panel). When these preconditions were met, the spheres and the fixation cross disappeared and after a SOA of 250 ms the target digit appeared at the center of one of the 10 (first experiment) or four (second experiment) displacement indicators (see **Figure 4**, right panel). Red, 3D mesh models of Arabic single-digits (1–9, except 5) were used as target stimuli. Digits were 7.7 cm in height and subtended a visual angle of 19.5°, 15.20°, 12.45°, 10.55°, 9.15°, 8.08°, 7.23°, 6.54°, 5.97°, and 5.50° at the different presentation distances, respectively.

Trials were canceled if the response took longer than 2000 ms. Furthermore, trials were canceled if the hands left the initial position or if either hand was clenched during the 250 ms SOA between the offset of the fixation cross and the presentation of the target stimulus (0.35%/0.38% of all trials in the first experiment/second experiment). The respective trials were repeated at the end of the block. In case of time-outs (more than 2000 ms), early movements (less than 250 ms, that is, within the SOA), or wrong responses, participants received according feedback. If the response was correct, participants received positive feedback. The whole experiment was self-paced, since trials only started when participants took the initial position and fixated the fixation cross. Hence, participants could (and they were encouraged to) take breaks between trials at any time, but they were not allowed to take off the HMD during breaks. All participants tolerated the VR procedure well and no experimental session was canceled.

Participants could practice the magnitude judgment before the actual blocks. In these training trials, participants responded with their left hand in case of a small (1) and with their right hand in case of a large practice digit (10); note that the large practice digit was not part of the actual stimulus set during testing. After completing 20 trials correctly, participants were allowed to proceed with the actual blocks.

Factors, Measures, Data Treatment

In both experiments, we varied two factors across trials and one factor across blocks. First, the *spatial displacement* of the target digit in the radial axis varied. In the first experiment, 10 equally-spaced radial displacements were used. The physical distance between two adjacent displacement indicators was about 10 cm (see **Figure 1**). In the second experiment, only four out of the 10 initial displacements were used; here, the physical distance between two adjacent distance indicators was about 20 cm (yellow lines in **Figure 1**). Second, the *digit magnitude* varied, we used the digits from 4 to 4 and 6 to 9 as target stimuli. Third, the response mapping varied between blocks, in one block participants responded with the left/right hand to small/large stimuli, in the other block, this mapping was reversed. In the analysis, this factor was recoded as *response hand* – either left or right. Each of the 80 (first experiment) or 32 (second experiment) displacement × digit combinations were repeated 4 (first experiment) or 10 (second experiment) times per block, yielding 320 trials per block. Trial and block

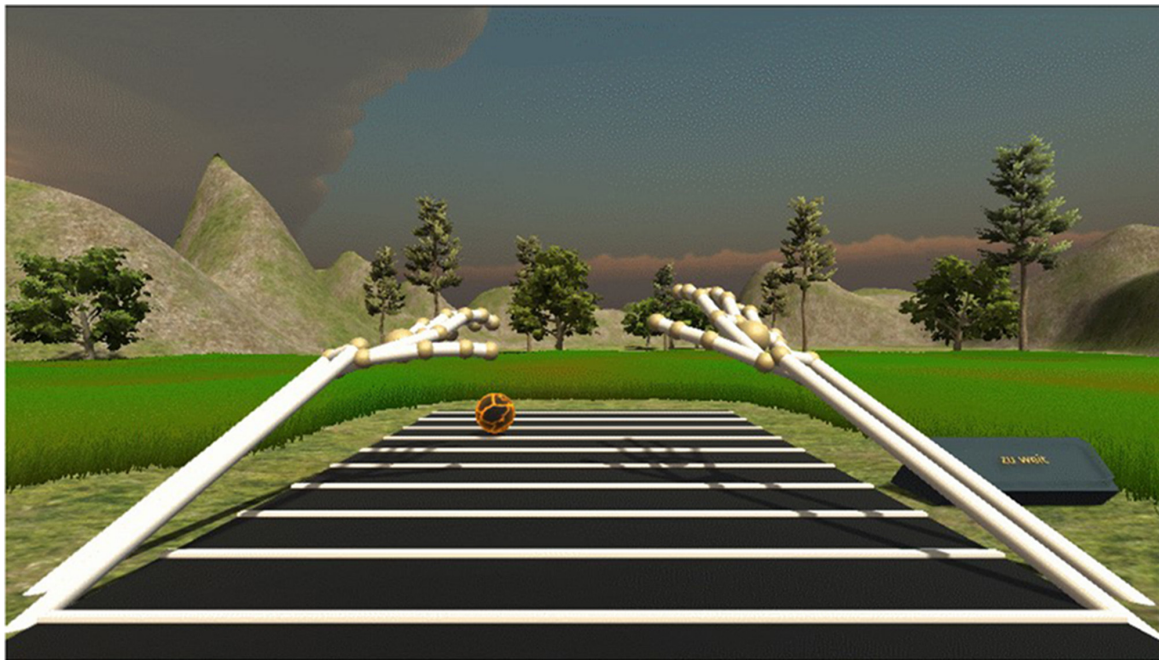


FIGURE 3 | The sensorimotor exploration task. Colored spheres appeared at different displacements and participants were requested to touch them with the correct hand (yellow spheres = left hand, green spheres = right hand). If a sphere appeared at an unreachable distance, participants were requested to push the button on the right side, labeled as “too far” (“zu weit,” in German). This task was intended to familiarize the participants with the VR environment and to provide a behavioral experience of reachability and distance.

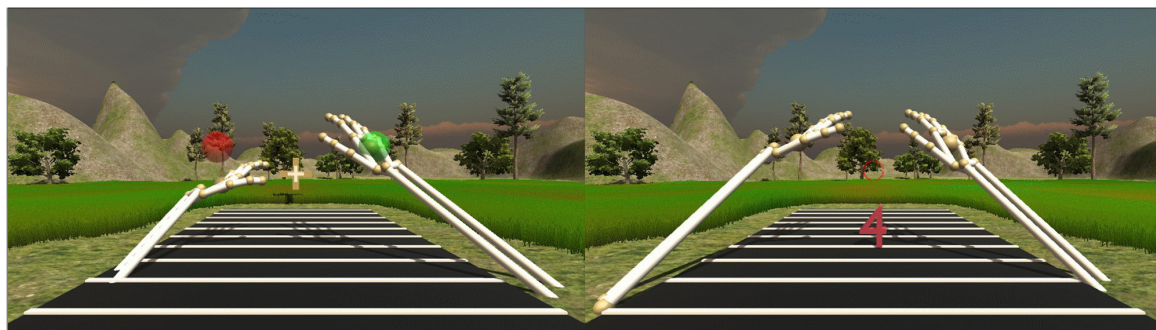


FIGURE 4 | The magnitude judgment task. Preconditions (**left**): open palms had to be placed correctly into initial positions (spheres turning green) and head rotation had to focus the outer-bound fixation cross for 2000 ms. Trial (**right**): a single-digit target at one of the displacements had to be classified as being smaller or larger than 5. Participants had to respond by clenching their fist as fast as possible while keeping their hands at the initial positions.

order was randomized. We recorded correct response times (RTs) in the magnitude judgment task and computed medians for all factor combinations. Furthermore, we recorded the maximum hand closure (MHC) of the irrelevant (incorrect) hand in each trial, as well as the respective time of the maximum hand closure time (MHCT). The MHC measure was thought to roughly reflect the degree of involuntary response preparation amid eventually correct responding especially for incongruent trials. Data from error trials were excluded from the analyses (4.2% in the first experiment and 4.7% in the second experiment). Before the analysis, RT outliers above or below two standard deviations from the respective cell mean were excluded as

well (0.2% in the first experiment¹ and 3.8% in the second experiment).

RESULTS

Seeing that the first and second experiment only differed regarding the number of spatial displacements, to focus the

¹Each cell mean was obtained from four data points, hence nearly no data points were excluded. The small number of repetitions per condition was the main motivation to rely on median RTs, as the median provides a less biased estimate in case of few observations.

analysis, and to increase the statistical power, we here report the results from the combined analysis with the between factor *experiment* for all $N = 32$ participants, considering only the four displacements applied in both experiments (close to the body, close to the hands, at the border between peripersonal and extrapersonal space, and in extrapersonal space). To anticipate, the between-experiments factor was not significant in any analysis and results were overall comparable. We report repeated measures ANOVAs and regression coefficient analyses based on RTs, MHC, and MHCT data. All ANOVAs were carried out with type III Sums of Squares. In case of violations of the assumption of sphericity, the respective p -values were submitted to a Greenhouse–Geisser correction. All p -values obtained from *post hoc* t -tests were submitted to a Bonferroni–Holm adjustment to correct for multiple comparisons. Data from the IPQ questionnaires was compared with reference data using independent sample t -tests regarding the three scales spatial presence, involvement, and realism.

Response Times

The repeated-measures ANOVA on median RTs joined from both experiments yielded a significant main effect of *digit magnitude* [$F(7,210) = 15.04, p < 0.001, \eta_p^2 = 0.33$]. The two-way interaction between *digit magnitude* \times *response hand* [$F(7,210) = 23.17, p < 0.001, \eta_p^2 = 0.44$] was significant as well. Further inspection of the main effect of *digit magnitude* revealed a numerical distance effect in terms of slower responses to digits 4 (606 ms) and 6 (620 ms), respectively, compared to 1 [564 ms; $t(31) = 6.65, p < 0.001$] and to 9 [569 ms; $t(31) = 6.93, p < 0.001$]. The two-way interaction effect between *response hand* \times *digit magnitude* indicated the typical horizontal SNARC effect: judgments for relatively small digits (less than 5) were faster performed with the left hand (552 ms) than with the right hand [616 ms; $t(31) = 6.17, p < 0.001$]. Vice versa, responses for large digits (greater than 5) were faster for the right hand (558 ms) than for the left hand [624 ms; $t(31) = 4.87, p < 0.001$].

There was no indication of a radial SNARC effect in the radial viewing dimension in the two-way interaction between *spatial displacement* \times *response hand* [$F(3,90) = 0.27, p = 0.812, \eta_p^2 = 0.01$]. Furthermore, the three-way interaction for *digit magnitude* \times *spatial displacement* \times *response hand* was not significant [$F(21,630) = 1.18, p = 0.308, \eta_p^2 = 0.04$].

In line with generally comparable data sets, the between-subjects main effect *experiment* was not significant [$F(1,30) = 1.68, p = 0.205, \eta_p^2 = 0.05$]. Importantly, both the four-way interaction between *experiment* \times *spatial displacement* \times *digit magnitude* \times *response hand* [$F(21,630) = 0.81, p = 0.609, \eta_p^2 = 0.03$] and the three-way interaction between *experiment* \times *digit magnitude* \times *response hand* [$F(7,210) = 1.08, p = 0.359, \eta_p^2 = 0.03$] were not significant as well, suggesting comparable SNARC effects for the two data sets. However, there was a trending two-way interaction between *response hand* \times *experiment* [$F(1,30) = 3.10, p = 0.088, \eta_p^2 = 0.09$] and participants in the first experiment were in general somewhat faster for right-hand responses (mean dRT = -9.7 ms), opposite to the behavior of participants in the second experiment (mean dRT = 8.5 ms). Finally, the ANOVA also revealed a trending main

effect of *spatial displacement* [$F(3,90) = 2.90, p = 0.057, \eta_p^2 = 0.09$]: participants were fastest if target stimuli were presented at the border of peripersonal space (580 ms) as compared to the displacements close to the body [587 ms; $t(31) = 1.83, p = 0.153$], close to the hands [591 ms; $t(31) = 2.30, p = 0.071$], and compared to the presentation in extrapersonal space [591 ms; $t(31) = 3.06, p = 0.027$].

We next inspected the modulation of horizontal SNARC effects by the visual presentation of targets at the different spatial displacements. Following the standard linear regression procedure for assessing SNARC effects (Lorch and Myers, 1990; Fias et al., 1996), we separately extracted for each participant and each of the four spatial displacements the correlation coefficient between numerical magnitude and response hand RT difference (dRT = right hand RT – left hand RT). More precisely, in this regression coefficient analysis, the response hand RT differences are predicted by the numerical magnitude factor (1, 2, 3, 4, 6, 7, 8, 9). Negative coefficients are indicative of relatively faster left-hand responses to smaller digits and of relatively faster right-hand responses to larger digits, which realizes the substantial result of the horizontal SNARC effect (see Figure 5).

Throughout both studies and across all four *spatial displacements*, the regression coefficient analysis yielded negative signed coefficients, as expected for horizontal SNARC effects (means and test statistics are reported in Table 1, data are shown in Figures 5, 6). All extracted coefficients were submitted to a mixed ANOVA comprising the repeated measures factors *spatial displacement* and the group variable *experiment*. The analysis yielded a significant main effect of *spatial displacement* [$F(3,90) = 3.60, p = 0.026, \eta_p^2 = 0.11$]. The two-way interaction of *spatial displacement* \times *experiment* was not significant [$F(3,90) = 0.83, p = 0.481$], and we neither observed a simple main effect of *experiment* [$F(1,30) = 0.04, p = 0.836$].

In general, the results show a relatively complex modulation of horizontal SNARC effects by spatial displacement (cf. Table 1). Paired t -tests were performed to compare SNARC effects for the different displacements. The SNARC effect close to the hands was significantly larger than the SNARC effect in the border-condition [$t(31) = -3.16, p = 0.012$], and tended to be larger than the SNARC effect in the close-to-body condition [$t(31) = -1.88, p = 0.082$]. Furthermore, the border-condition SNARC effect tended to be smaller than the SNARC effect in extrapersonal space [$t(31) = 2.14, p = 0.082$]. All remaining comparisons were statistically not significant ($ts < 1.55$).

Maximum Hand Closure (MHC) and Maximum Hand Closure Time (MHCT)

Based on the response-related conflict elicited by SNARC effects in different previous EEG studies (e.g., Keus et al., 2005), and previous results on number-action links (e.g., Andres et al., 2004), we expected to observe a tendency for spatial-numerical associations also in the continuous activation of responding (i.e., closing the hand) in the SNARC-congruent, yet false response (i.e., in the incongruent block as opposed to the congruent block). To inspect this potential behavior, the continuous closure of the incorrect hand during correct responding was recorded in the

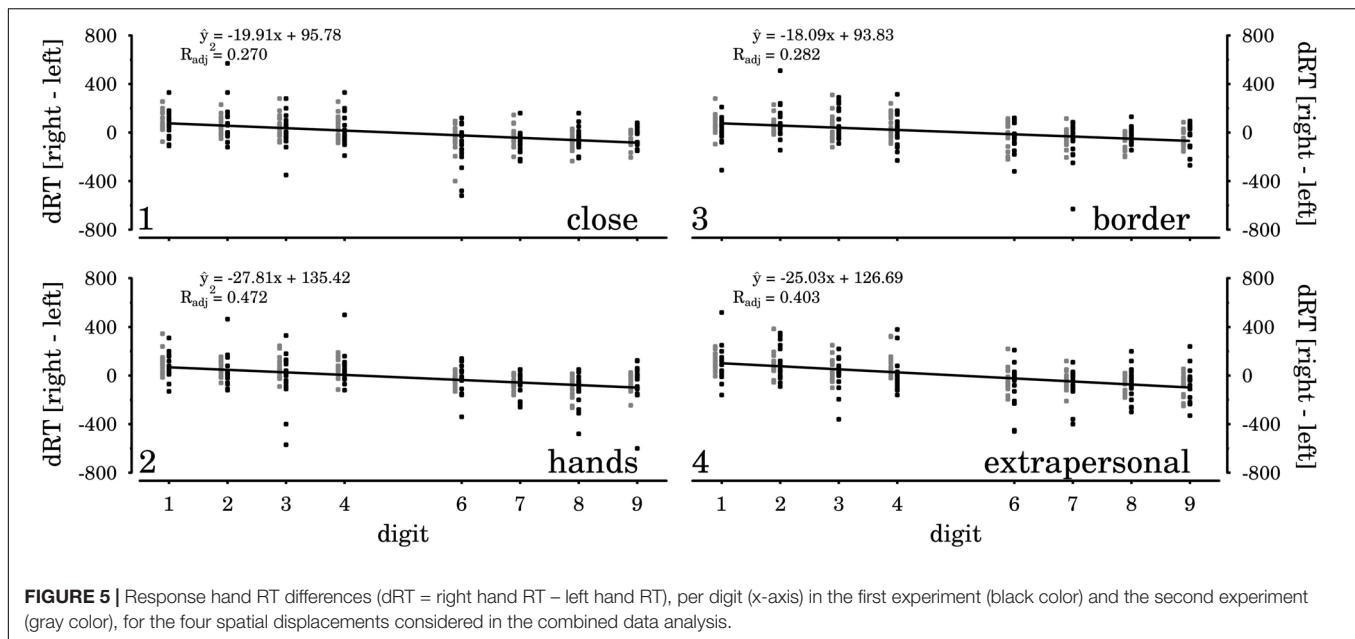


FIGURE 5 | Response hand RT differences (dRT = right hand RT – left hand RT), per digit (x-axis) in the first experiment (black color) and the second experiment (gray color), for the four spatial displacements considered in the combined data analysis.

TABLE 1 | Horizontal SNARC effects resulting from the regression coefficient analysis for both studies at the four considered displacements (means and standard deviations in ms/magnitude bin and in hand closure unit/magnitude bin).

Displacement	RT		MHCT		MHC	
	M	SD	M	SD	M	SD
Four extrapersonal space	-25.0**	23.9	-19.6**	24.8	-0.0033*	0.0080
Three border	-18.1**	19.6	-8.4*	23.1	-0.0024*	0.0069
Two close to hand	-27.8**	28.9	-25.3**	27.7	-0.0023*	0.0049
One close to body	-19.9**	17.7	-10.9**	23.4	-0.0018	0.0060

Asterisks indicate that coefficients differ significantly from zero. * $p < 0.05$; ** $p < 0.005$ (one-tailed).

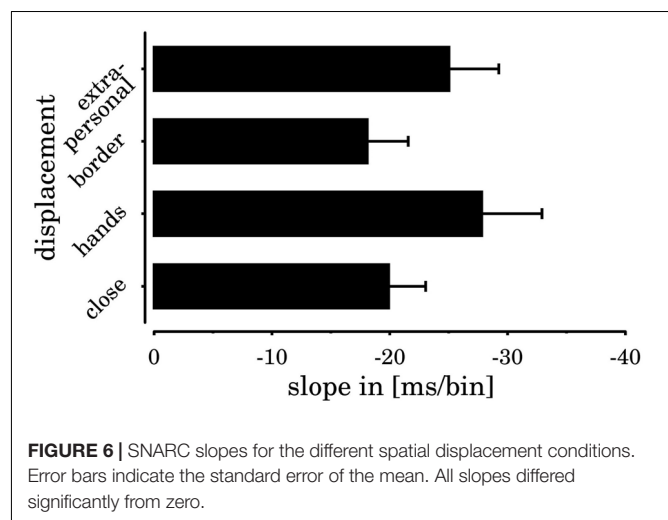
VR framework as dependent variable². Based on this trajectory, we obtained the MHC per trial, as well as the according time (MHCT), relative to target onset.

For approximately one quarter of all participants ($N_{\text{First}} = 3$, $N_{\text{Second}} = 4$), this sort of analysis was not possible because these participants kept their incorrect hands perfectly open during responding and thus the value was continuously zero. For the remaining $N = 25$ participants³, data were submitted to the ANOVA and regression coefficient analysis as before, using MHC and MHCT as dependent variables.

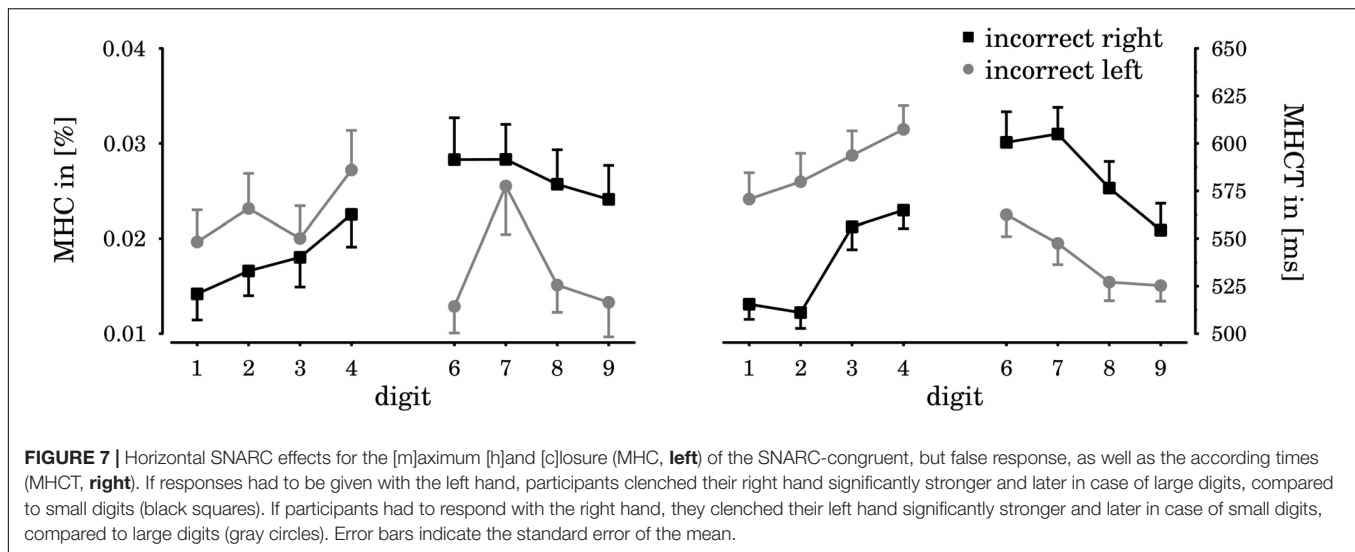
The ANOVA on MHC revealed a significant two-way interaction between *numerical magnitude* \times *response hand* [$F(7,161) = 2.92$, $p = 0.007$, $\eta_p^2 = 0.11$; see also **Figure 7**]. *Post hoc t*-tests revealed a tendency for a horizontal SNARC effect. If participants had to respond to relatively large digits (greater than 5), incorrect closure of the right hand was stronger (0.027)

²This value equals 0 if the hand is open, a value of 1 indicates a clenched fist. The programming interface of the LeapMotion® sensor allows to record this measure directly.

³Unfortunately, reduced and unequal sample sizes regarding the between factor experiment ($N_{\text{First}} = 13$, $N_{\text{Second}} = 12$) resulted from this procedure. The applied type III sums of square provide an adequate adjustment, but results may still be relatively insensitive regarding possible smaller group differences.



than incorrect closure of the left hand [0.017; $t(24) = 2.14$, $p = 0.043$]. In case of relatively small digits (less than 5), the respective difference was not significant [$t(24) = 0.81$, $p = 0.428$], but the overall pattern of results fitted a typical



horizontal SNARC effect (see **Figure 7**). As before, there was no significant indication of a radial SNARC effect and the two-way interaction between *spatial displacement* \times *response hand* was not statistically significant [$F(3,69) = 1.67, p = 0.181, \eta_p^2 = 0.07$]. The three-way interaction between *numerical magnitude* \times *spatial displacement* \times *response hand* was not statistically significant, either [$F(21,483) = 1.43, p = 0.182, \eta_p^2 = 0.06$], as well as the main effect for *numerical magnitude* [$F(7,161) = 1.80, p = 0.146, \eta_p^2 = 0.07$]. There was no main effect for the between factor experiment [$F(1,23) = 0.87, p = 0.361, \eta_p^2 = 0.04$], and no interactions involving this factor reached significance ($ps > 0.12$).

As with RTs, we also performed regression coefficient analysis and obtained consistently negative-signed coefficients (see **Table 1**). In contrast to the RT analyses, the differences between coefficients on MHC were not significant ($ts < 0.88, ps > 0.39$). Furthermore, t -tests against zero detected only a trending significance for the negative-signed regression coefficient in the condition close to the body [$t(24) = 1.50, p = 0.074$; one-tailed]. Horizontal SNARC effects themselves were significantly smaller than zero in all remaining conditions [close to the hands: $t(24) = 2.36, p = 0.014$; at border: $t(24) = 1.72, p = 0.049$; in extrapersonal space: $t(24) = 2.05, p = 0.026$].

Regarding MHCT, the ANOVA revealed significant main effects of *digit magnitude* [$F(7,161) = 8.31, p < 0.001, \eta_p^2 = 0.27$] and *spatial displacement* [$F(3,69) = 2.77, p = 0.048, \eta_p^2 = 0.11$]. The two-way interaction between *digit magnitude* \times *response hand* [$F(7,161) = 7.63, p < 0.001, \eta_p^2 = 0.25$] was significant as well. All remaining effects did not reach significance ($ps > 0.10$). *Post hoc* t -tests for the main effect of *digit magnitude* revealed slower responses to digits four (586 ms) and six (581 ms), respectively, compared to one [543 ms; $t(24) = 6.26, p < 0.001$] and to nine [539 ms; $t(24) = 5.51, p < 0.001$], thus mimicking the numerical distance effect. Further analysis of the two-way interaction effect between *response hand* \times *digit magnitude* revealed a horizontal SNARC effect (see **Figure 7**): in case of relatively small digits (less than 5), MHCT occurred earlier for

left hand responses (536 ms) as compared to right hand responses [587 ms; $t(24) = 3.53, p < 0.01$]. Vice versa, MHCT in case of relatively large digits (greater than 5) occurred earlier for right hand responses (540 ms) than for left hand responses [584 ms; $t(24) = 3.34, p < 0.01$]. *Post hoc* analysis of the *spatial displacement* main effect showed that MHCT occurred earlier for stimuli presented at the border of peripersonal space (554 ms), compared to stimuli presented close to the hands [570 ms; $t(24) = 3.03, p = 0.018$]. The comparison with stimuli presented close to the body [564 ms; $t(24) = 1.56, p = 0.266$] and in extrapersonal space [560 ms; $t(24) = 0.79, p = 0.439$] yielded no significant differences. In general, the pattern of results obtained in MHCT was similar to the observed pattern in RT. Indeed, both measures were highly correlated [$r(1598) = 0.74, p < 0.001$]. On average, MHCT occurred only shortly before the actual response [$M_{MHCT-RT} = -29$ ms, $SD_{MHCT-RT} = 32$ ms; $t(24) = 4.50, p < 0.01$].

An analysis of the regression coefficients obtained from the MHCTs yielded a significant main effect of *spatial displacement* [$F(3,69) = 4.48, p = 0.006, \eta_p^2 = 0.16$]. The slopes for stimuli presented close to the hands were more inclined than slopes in the border-condition [$t(24) = -3.47, p = 0.024$] and in the close-to-body condition [$t(24) = -3.01, p = 0.030$]. Again, these results dovetail with the RT pattern (cf. **Table 1**). There were no effects of the group variable *experiment* ($ps > 0.31$).

IPQ Data

Self-reported ratings of presence (IPQ questionnaire) obtained from the 32 participants were compared with reference data provided by the igroup consortium (see **Table 2**⁴). The reference data set was obtained from video games where the players were equipped with an HMD and comprised 24 mean values for the three subscales. Independent sample t -tests yielded a significant difference for spatial presence [$t(31.31) = 2.08, p = 0.022$]. Compared to the reference data, participants in our setup

⁴<http://www.igroup.org/pq/ipq/data.php>

TABLE 2 | Self-report ratings of presence (IPQ questionnaire).

IPQ presence component	Observed data		Reference data	
	<i>M</i>	<i>SEM</i>	<i>M</i>	<i>SEM</i>
Spatial presence	4.18	0.13	3.46	0.32
Involvement	2.51	0.20	2.59	0.23
Realism	2.30	0.14	2.06	0.26
Mean score	2.99	0.13	2.70	0.17

Ratings range on seven-point Likert scale from -3 (not at all) to $+3$ (very much). For the analysis, the value range was recoded to fit a scale from 0 to 6, in accordance with the evaluation guidelines proposed by the iGroup consortium.

reported a higher degree of spatial presence. With respect to involvement and realism, our data compares to the reference ($ps > 0.206$). Together, the results show a sufficient degree of immersion. Improvements with respect to spatial presence dovetail with our earlier results obtained in setups where we applied the LeapMotion[®] sensor together with an Oculus Rift[®] DK2 HMD (Schroeder et al., 2016; Lohmann and Butz, 2017).

To evaluate correlations between the horizontal SNARC effects at different spatial displacements with the subjective presence experience in the virtual environment, Pearson correlation coefficients were computed for each IPQ subscale. Relatively high coefficients were obtained for the spatial presence subscale at all four spatial displacements [$r(31) = 0.17$ to $r(31) = 0.33$]. For stimuli presented close to the hands, the correlation was most pronounced [$r(31) = 0.331$, $p = 0.064$]. The correlations between SNARC effect (as obtained in the regression coefficient analysis) and involvement ($|r(31)| < 0.12$) and realism [$r(31) < 0.19$] were less pronounced.

DISCUSSION

In two experiments, we investigated effects of radial distance on numerical magnitude comparisons in an immersive VR. Results show a consistent, but complex pattern of interactions between spatial displacements, numerical magnitude, and side of responding: a horizontal SNARC effect was observed in terms of faster left-hand responses to relatively small digits and faster right-hand responses for relatively large digits. In kinematic parameters, we also observed the horizontal SNARC effect in terms of response activation in the incorrect hand (MHC) particularly for incongruent trials, shortly before the actual response (MHCT). Regarding the regression analyses, horizontal SNARC effects were most pronounced when target digits were presented close to the hands or in extrapersonal space, compared to other spatial displacements (close to the body or at the border of peripersonal space). Together, these results are in line with the assumption of a situated, sensorimotor representation underlying spatial-numerical associations that supports flexible spatial-numerical mappings.

The Relationship Between Reachability and SNARC Effects

The results show a robust horizontal SNARC effect for all tested spatial displacements. However, the pattern of results

is inconsistent with a linear relationship between numerical magnitude, response side, and physical distance. Instead, regression coefficient analyses revealed that SNARC slopes were most inclined when stimuli were presented near the hands or just outside reachable space. The pronounced SNARC slopes near the hands seem not to be due to a mere near hand effect (Reed et al., 2006), which would predict faster RTs for stimuli presented near the hand in general. Instead, the steeper slopes in this spatial displacement are actually in line with the notion that spatial attention is more specifically subject to altered cognitive processing when objects approach the own hands (Abrams et al., 2008; Tseng et al., 2012). For example, it has been shown that the processing of stimuli close to the hands involves both costs, like delayed disengagement, and benefits, for instance reduced distraction by task-irrelevant features (Davoli and Brockmole, 2012; Liepelt and Fischer, 2016).

In the extrapersonal condition, horizontal spatial-numerical associations were present. This observation may be considered to be in conflict with studies showing that object affordances are limited to peri-personal space (e.g., Costantini et al., 2011; Kalénine et al., 2016). For instance, counting direction preference was reduced when children interacted with counting objects in extrapersonal space using a laser pointer (Patro et al., 2015). However, it is important to emphasize that number presentation within or outside of peripersonal space was task-irrelevant in our study and participants did not perform grasp movements, but classified the presented numbers as being small or large by adjacent left-hand or right-hand closure without further movement, functionally rendering object affordances meaningless for correct responding. Given that Andres et al. (2004) observed an association between grip closure with small numbers, it is still conceivable that using hand closure as response mode in the present experiments induced overall biases in favor of small numbers (and perhaps left space).

So far, there have been no studies that investigated the effects and interactions of different spatial directional codes in the visual presentation dimension on the SNARC effect. If magnitudes in different modalities are mapped directly, one would expect a linear relation between spatial magnitude, e.g., radial distance, and numerical magnitude, which would yield a radial SNARC effect. Our results provide no evidence for such an effect, extending the findings of two earlier studies. Santens and Gevers (2008) had their participants respond to large or small digits with either close or far movements. Close responses were faster for small digits, whereas far responses were relatively faster for large digits. Marghetis and Youngstrom (2014) found evidence for a radial SNARC effect when they let participants respond to positive and negative integers by stepping forward or backward. Here, forward movements yielded faster RTs in case of positive integers, while responses for negative integers were faster in case of backward movements. This compatibility effect vanished when only positive integers were used as stimuli. In both studies, the spatial displacement of the target digits was not manipulated. The results show a semantic overlap between numerical magnitude and response distance (Santens and Gevers, 2008) or between positive- and negative- numbers and response direction (Marghetis and Youngstrom, 2014),

respectively. Hence, both results do not necessarily imply a relation between radial distance and SNARC magnitude, but between movement magnitude and SNARC magnitude, that is, a number-action instead of a number-space link. The assumed dominance of a number-action link would also provide an explanation why the effect size of the interactions between horizontal response dimension and the semantic magnitude, which were both task-relevant, are much larger when compared to any effect of the perceptual magnitude in the radial dimension, which was task-irrelevant in both experiments we reported here. Although it was long assumed that numerical magnitude biases cognitive processing automatically, some previous results actually show that very basic perceptual decision tasks can tremendously diminish the influence of spatial-numerical processing (Fias et al., 2001; Schroeder et al., 2017). Moreover, associations between numerical magnitude and radial distance were observed in tasks that positioned effectors accordingly along the distance dimension (Müller and Schwarz, 2007; Gronau et al., 2017).

Theoretically, these results also further specify the taxonomy of spatial-numerical associations, which pits the implicit directional effects as observed in SNARC tasks against other explicit linkages and non-directional links between space and number (Patro et al., 2014; Cipora et al., 2015). We propose that the exact spatial direction of number mappings is determined by situated and task-relevant implementation of action, i.e., using left-hand and right-hand responding, rather than low-level processing of irrelevant spatial information. In the virtual environments, visual cues of depth information (e.g., in terms of number symbol size) further emphasized this type of magnitude information, as closer numbers were larger. However, even this salient relation between numerical magnitude, number symbol size, and distance did not yield interactions reflecting a direct mapping between these types of magnitude information⁵. VR allows to disentangle contributions of these different dimensions and future studies can systematically test this prediction of different situated conditions, which contrasts with previous accounts of generally weaker or steeper vertical SNARC effects. In line with earlier findings (e.g., Andres et al., 2004; Lindemann et al., 2007), these results imply a relationship between action parameters and numerical cognition, which indicates that spatial-numerical associations are realized within a sensorimotor metric. This interpretation is further corroborated by the observed correlation between kinematic parameters (MHC) and numerical magnitude.

Response Conflict and Effects on Kinematic Parameters

Different experiments have provided evidence for the assumption that the SNARC effect arises at a late, response-related stage of processing (e.g., Keus et al., 2005). For instance, robust SNARC effects have been observed in response-locked

event-related potentials (ERPs), while they were absent at earlier ERPs, associated with stimulus processing (Fischer and Miller, 2008). Furthermore, Vierck and Kiesel (2010) showed a compatibility effect between response force and numerical magnitude. Participants responded faster when small digits required a weak response force – while for large digits the opposite was true. Complementary to these findings, our mean hand closure measurements show relatively consistent activations of the incorrect, but SNARC-compatible responses in case of SNARC-incompatible response mappings. Specifically, if participants had to respond to large digits with their left hand, they clenched their right hand significantly stronger than when they had to respond to small digits with their left hand, and vice versa. The temporal pattern of this clenching was highly similar to the RT pattern and showed a similar modulation by spatial displacement. Apparently, response selection was primed by numerical magnitude (see also Daar and Pratt, 2008). Even considering the large interindividual difference in the extent of the SNARC effect (e.g., Wood et al., 2008) and also the amount of response preparation in incorrect hands, which was not reliably available for analysis in one quarter of our participants and also relatively weak (the maximum observed value was ~ 0.20 , but the threshold for responding was 0.75; see **Figure 7**), this pattern of results again implies a strong grounding of spatial-numerical mappings in a sensorimotor metric. The SNARC slopes for the hand closure did not change with physical distance, however, slopes for the hand closure time showed the same systematicities as slopes obtained from RTs. Given the small sample size, interindividual differences, and the resistance against involuntary hand movements in a forth of our population, it remains open whether response execution was unaffected by reachability. Furthermore, more extensive responses may be better suited to yield variable measures and to detect more subtle interaction terms in future research, as it was also recommended in mouse tracking research (Fischer and Hartmann, 2014; Pinheiro-Chagas et al., 2017).

Regarding the size of effects of numerical magnitude on kinematic parameters, we only observed an effect on the horizontal SNARC effect and apparently the measurement as well as assessment in the VR parameter space was a little noisier, at least as compared to RT assessments. This might indicate a certain specificity of the response parameters. The applied VR setup allowed a convenient manipulation of perceived physical distance and perceived reachability of the target stimuli. Furthermore, motion capture allowed to record a continuous response and to detect SNARC effects within the kinematic parameters of response execution. In general, VR setups seem well-suited to further investigate the role of sensorimotor codes in numerical cognition – especially with respect to different spatial mappings of stimuli and effectors. In future studies, the establishment of the VR procedure and validation for horizontal SNARC effects in the current study allows for further perceptual as well as bodily manipulations, such as the manipulation of the perceived reachability by manipulating the virtual arm length or by the induction of multisensory conflict (e.g., Lohmann and Butz, 2017). As it was argued by Viarouge et al. (2014) spatial-numeric mappings can be

⁵The apparent absence of effects of physical stimulus size fits well with the results of Longo and Lourenco (2010), who showed that the observed effects of perceptual distance on line-bisection and number-bisection were not affected by changes in physical stimulus size.

established on the fly within different frames of reference, depending on the experimental context, like instructions or saliency of spatial anchors. This impact of situated influences on spatial-numerical mappings led to the formulation of a taxonomy to structure these influences (Cipora et al., 2016). The outlined VR paradigms and manipulations of sensorimotor mappings as well as spatial perception will allow a more detailed investigation of the contextual parameters that give rise to certain spatial-numerical mappings, and to clarify the effect of action-related manipulations on spatial-numerical mappings. For instance, the present results imply that SNARC effects are bound to the response relevant spatial axis, instead of a general dominance of either the horizontal, or the vertical axes.

CONCLUSION

Although interactions between semantic and perceptual magnitudes are well-known (e.g., Henik and Tzelgov, 1982; Cohen-Kadosh et al., 2008), the exact shape of these interactions is not clear and corresponding theories were often underspecified, i.e., by generalizing the common code to all possible dimensions. Our results imply that spatial-numeric mappings between different magnitude codes are constrained by task-relevance and characteristics of the sensorimotor metric in which they are realized.

We did not observe interactions between task-relevant horizontal responses and task-irrelevant radial spatial displacements. However, the standard horizontal SNARC effect between task-relevant horizontal responses and task-relevant semantic magnitude was convincingly demonstrated in the immersive VR and further transferred in hand closure measurements to show response competition in the non-responding hand. The systematic manipulation of spatial displacements in the stereoscopic display furthermore revealed a non-linear interaction between physical distance and SNARC magnitude. Given these findings, it seems highly likely that

spatial-numerical associations are driven by a sensorimotor metric, which is situated on the fly in the current task-demands. The selective emphasis of action-relevant processing close to effectors is generally consistent with both, theories of anticipatory behavior control and with the parietal foundations of action-relevant numerical processing. The apparent complex interactions, however – particularly when presentations exceeded the peripersonal perceptual space – call for further systematic explorations and theoretical considerations of body-related cognitive processing. Furthermore, the observed tendency for a relation between spatial presence and magnitude of the SNARC effect requires further investigation.

ETHICS STATEMENT

All participants volunteered and provided written informed consent. The study was conducted in accordance with German Psychological Society (DGPs) ethical guidelines (2004, CIII), which are in accordance with the WMA declaration of Helsinki.

AUTHOR CONTRIBUTIONS

JL, PS, H-CN, CP, and MB developed the study concept. PS and JL designed the study, performed the data analysis and interpretation, and drafted the manuscript. JL implemented the study, performed the testing, and data collection. H-CN, CP, and MB provided critical revisions. All authors approved the final version of the manuscript for submission.

FUNDING

We acknowledge the support by Deutsche Forschungsgemeinschaft and the Open Access Publishing Fund of the University of Tübingen.

REFERENCES

- Abrams, R. A., Davoli, C. C., Du, F., Knapp, W. H., and Paull, D. (2008). Altered vision near the hands. *Cognition* 107, 1035–1047. doi: 10.1016/j.cognition.2007.09.006
- Andres, M., Davare, M., Pesenti, M., Olivier, E., and Seron, X. (2004). Number magnitude and grip aperture interaction. *Neuroreport* 15, 2773–2777.
- Bueti, D., and Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1831–1840. doi: 10.1098/rstb.2009.0028
- Cipora, K., Patro, K., and Nuerk, H.-C. (2015). Are spatial-numerical associations a cornerstone for arithmetic learning? The lack of genuine correlations suggests: no. *Mind Brain Educ.* 9, 190–207. doi: 10.1111/mbe.12093
- Cipora, K., Patro, K., and Nuerk, H.-C. (2016). “Situated influences on spatial-numerical associations,” in *Spatial Biases in Perception and Cognition*, ed. T. Hubbard (Cambridge: Cambridge University Press).
- Cohen-Kadosh, R., and Dowker, A. (eds) (2015). *The Oxford Handbook of Numerical Cognition*. Oxford: Oxford Library of Psychology.
- Cohen-Kadosh, R., Lammertyn, J., and Izard, V. (2008). Are numbers special? An overview of chronometric, neuroimaging, developmental and comparative studies of magnitude representation. *Prog. Neurobiol.* 84, 132–147. doi: 10.1016/j.pneurobio.2007.11.001
- Costantini, M., Ambrosini, E., Scorolli, C., and Borghi, A. M. (2011). When objects are close to me: affordances in the peripersonal space. *Psychon. Bull. Rev.* 18, 302–308. doi: 10.3758/s13423-011-0054-4
- Daar, M., and Pratt, J. (2008). Digits affect actions: the SNARC effect and response selection. *Cortex* 44, 400–405. doi: 10.1016/j.cortex.2007.12.003
- Dackermann, T., Fischer, U., Nuerk, H., and Cress, U. (2017). Applying embodied cognition: from useful interventions and their theoretical underpinnings to practical applications. *ZDM* 49, 545–557. doi: 10.1007/s11858-017-0850-z
- Davoli, C. C., and Brockmole, J. R. (2012). The hands shield attention from visual interference. *Atten. Percept. Psychophys.* 74, 1386–1390. doi: 10.3758/s13414-012-0351-7
- Dehaene, S., Bossini, S., and Giraux, P. (1993). The mental representation of parity and number magnitude. *J. Exp. Psychol. Gen.* 122, 371–396. doi: 10.1037/0096-3445.122.3.371
- Dehaene, S., Piazza, M., Pinel, P., and Cohen, L. (2003). Three parietal circuits for number processing. *Cogn. Neuropsychol.* 20, 487–506. doi: 10.1080/02643290244000239
- Fairhurst, M., and Deroy, O. (2017). Testing the shared spatial representation of magnitude of auditory and visual intensity. *J. Exp. Psychol. Hum. Percept. Perform.* 43, 629–637. doi: 10.1037/xhp0000332

- Fias, W., Brysbaert, M., Geypens, F., and D'Ydewalle, G. (1996). The importance of magnitude information in numerical processing: evidence from the SNARC effect. *Math. Cogn.* 2, 95–110. doi: 10.1080/135467996387552
- Fias, W., Lauwereyns, J., and Lammertyn, J. (2001). Irrelevant digits affect feature-based attention depending on the overlap of neural circuits. *Cogn. Brain Res.* 12, 415–423. doi: 10.1016/S0926-6410(01)00078-7
- Fischer, M. H. (2005). Perceived reachability: the roles of handedness and hemifield. *Exp. Brain Res.* 160, 283–289. doi: 10.1007/s00221-004-2007-x
- Fischer, M. H., Castel, A. D., Dodd, M. D., and Pratt, J. (2003). Perceiving numbers causes spatial shifts of attention. *Nat. Neurosci.* 6, 555–556. doi: 10.1038/nn1066
- Fischer, M. H., and Hartmann, M. (2014). Pushing forward in embodied cognition: may we mouse the mathematical mind? *Front. Psychol.* 5:1315. doi: 10.3389/fpsyg.2014.01315
- Fischer, M. H., Riello, M., Giordano, B. L., and Rusconi, E. (2013). Singing numbers in cognitive space - a dual-task study of the link between pitch, space, and numbers. *Top. Cogn. Sci.* 5, 354–366. doi: 10.1111/tops.12017
- Fischer, R., and Miller, J. (2008). Does the semantic activation of quantity representations influence motor parameters? *Exp. Brain Res.* 189, 379–391. doi: 10.1007/s00221-008-1434-5
- Galton, F. (1880). Visualised numerals. *Nature* 21, 252–256. doi: 10.1038/021494e0
- Gamberini, L., Seraglia, B., and Priftis, K. (2008). Processing of peripersonal and extrapersonal space using tools: evidence from visual line bisection in real and virtual environments. *Neuropsychologia* 46, 1298–1304. doi: 10.1016/j.neuropsychologia.2007.12.016
- Gevers, W., Lammertyn, J., Notebaert, W., Verguts, T., and Fias, W. (2006). Automatic response activation of implicit spatial information: evidence from the SNARC effect. *Acta Psychol.* 122, 221–233. doi: 10.1016/j.actpsy.2005.11.004
- Gronau, N., Izoutcheev, A., Nave, T., and Henik, A. (2017). Counting distance: effects of egocentric distance on numerical perception. *PLoS One* 12:e0174772. doi: 10.1371/journal.pone.0174772
- Hartmann, M., Gashaj, V., Stahnke, A., and Mast, F. W. (2012). There is more than “More is up”: hand and foot responses reverse the vertical association of number magnitudes. *J. Exp. Psychol. Hum. Percept. Perform.* 40, 1416–1427. doi: 10.1037/a0026706
- Henik, A., and Tzelgov, J. (1982). Is three greater than five: the relation between physical and semantic size in comparison tasks. *Mem. Cogn.* 10, 389–395. doi: 10.3758/BF03202431
- Hesse, P. N., and Bremner, F. (2016). SNARC effect in different effectors. *Perception* 45, 180–195. doi: 10.1177/030100615614453
- Hoffmann, J. (1993). *Vorhersage und Erkenntnis: Die Funktion von Antizipationen in der Menschlichen Verhaltenssteuerung und Wahrnehmung. [Anticipation and cognition: The function of anticipations in human behavioral control and perception]*. Göttingen: Hogrefe.
- Hoffmann, J. (2003). “Anticipatory behavioral control,” in *Anticipatory Behavior in Adaptive Learning Systems: Foundations, Theories, and Systems*, eds M. V. Butz, O. Sigaud, and P. Gérard (Berlin: Springer-Verlag), 44–65. doi: 10.1007/978-3-540-45002-3_4
- Holmes, K. J., and Lourenco, S. F. (2011). Common spatial organization of number and emotional expression: a mental magnitude line. *Brain Cogn.* 77, 315–323. doi: 10.1016/j.bandc.2011.07.002
- Holmes, K. J., and Lourenco, S. F. (2012). Orienting numbers in mental space: horizontal organization trumps vertical. *Q. J. Exp. Psychol.* 65, 1044–1051. doi: 10.1080/17470218.2012.685079
- Holmes, N. P., and Spence, C. (2004). The body schema and multisensory representation (s) of peripersonal space. *Cogn. Process.* 5, 94–105. doi: 10.1007/s10339-004-0013-3
- Ito, Y., and Hatta, T. (2004). Spatial structure of quantitative representation of numbers: evidence from the SNARC effect. *Mem. Cogn.* 32, 662–673. doi: 10.3758/BF03195857
- Kalénine, S., Wamain, Y., Decroix, J., and Coello, Y. (2016). Conflict between object structural and functional affordances in peripersonal space. *Cognition* 155, 1–7. doi: 10.1016/j.cognition.2016.06.006
- Keus, I. M., Jenks, K. M., and Schwarz, W. (2005). Psychophysiological evidence that the SNARC effect has its functional locus in a response selection stage. *Cogn. Brain Res.* 24, 48–56. doi: 10.1016/j.cogbrainres.2004.12.005
- Liepert, R., and Fischer, R. (2016). Task demands determine hand posture bias on conflict processing in a Simon task. *Psychon. Bull. Rev.* 23, 579–586. doi: 10.3758/s13423-015-0901-9
- Lindemann, O., Abolafia, J. M., Girardi, G., and Bekkering, H. (2007). Getting a grip on numbers: numerical magnitude priming in object grasping. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 1400–1409. doi: 10.1037/0096-1523.33.6.1400
- Linkenauger, S. A., Bühlhoff, H. H., and Mohler, B. J. (2015). Virtual arm's reach influences perceived distances but only after experience reaching. *Neuropsychologia* 70, 393–401. doi: 10.1016/j.neuropsychologia.2014.10.034
- Lohmann, J., and Butz, M. V. (2017). Lost in space: multisensory conflict yields adaptation in spatial representations across frames of reference. *Cogn. Process.* 18, 211–228. doi: 10.1007/s10339-017-0798-5
- Longo, M. R., and Lourenco, S. F. (2010). Bisecting the mental number line in near and far space. *Brain Cogn.* 72, 362–367. doi: 10.1016/j.bandc.2009.10.016
- Lorch, R. F., and Myers, J. L. (1990). Regression analyses of repeated measures data in cognitive research. *J. Exp. Psychol. Learn. Mem. Cogn.* 16, 149–157. doi: 10.1037/0278-7393.16.1.149
- Marghetis, T., and Youngstrom, K. (2014). “Pierced by the number line: integers are associated with back - to - front sagittal space,” in *Proceedings of the 36th Annual Conference of the Cognitive Science Society*, eds P. Bello, M. Guarani, M. McShane, and B. Scassellati (Austin, TX: Cognitive Science Society), 946–951.
- Mohler, B. J., Bühlhoff, H. H., Thompson, W. B., and Creem-Regehr, S. H. (2008). “A full-body avatar improves egocentric distance judgments in an immersive virtual environment,” in *Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization*, APGV '08, (New York, NY: ACM), 194. doi: 10.1145/1394281.1394323
- Müller, D., and Schwarz, W. (2007). Is there an internal association of numbers to hands? The task set influences the nature of the SNARC effect. *Mem. Cogn.* 35, 1151–1161. doi: 10.3758/BF03193485
- Nuerk, H.-C., Wood, G., and Willmes, K. (2005). The universal SNARC effect: the association between number magnitude and space is amodal. *Exp. Psychol.* 52, 187–194. doi: 10.1027/1618-3169.52.3.187
- Patro, K., Nuerk, H.-C., and Cress, U. (2015). Does your body count? Embodied influences on the preferred counting direction of preschoolers. *J. Cogn. Psychol.* 27, 1–13. doi: 10.1080/20445911.2015.1008005
- Patro, K., Nuerk, H.-C., Cress, U., and Haman, M. (2014). How number-space relationships are assessed before formal schooling: a taxonomy proposal. *Front. Psychol.* 5:419. doi: 10.3389/fpsyg.2014.00419
- Pinheiro-Chagas, P., Dotan, D., Piazza, M., and Dehaene, S. (2017). Finger tracking reveals the covert stages of mental arithmetic. *Open Mind* 1, 30–41. doi: 10.1162/opmi
- Reed, C. L., Grubb, J. D., and Steele, C. (2006). Hands up: attentional prioritization of space near the hand. *J. Exp. Psychol. Hum. Percept. Perform.* 32, 166–177. doi: 10.1037/0096-1523.32.1.166
- Santens, S., and Gevers, W. (2008). The SNARC effect does not imply a mental number line. *Cognition* 108, 263–270. doi: 10.1016/j.cognition.2008.01.002
- Schroeder, P. A., Lohmann, J., Butz, M. V., and Plewnia, C. (2016). Behavioral bias for food reflected in hand movements: a preliminary study with healthy subjects. *Cyberpsychol. Behav. Soc. Netw.* 19, 120–126. doi: 10.1089/cyber.2015.0311
- Schroeder, P. A., Nuerk, H.-C., and Plewnia, C. (2017). Prefrontal neuromodulation reverses spatial associations of non-numerical sequences, but not numbers. *Biol. Psychol.* 128, 39–49. doi: 10.1016/j.biopsycho.2017.07.008
- Schroeder, P. A., and Pfister, R. (2015). Arbitrary numbers counter fair decisions: trails of markedness in card distribution. *Front. Psychol.* 6:240. doi: 10.3389/fpsyg.2015.00240
- Schubert, T. W., Friedmann, F., and Regenbrecht, H. (2001). The experience of presence: factor analytic insights. *Presence* 10, 266–281. doi: 10.1162/105474601300343603
- Schwarz, W., and Müller, D. (2006). Spatial associations in number-related tasks: a comparison of manual and pedal responses. *Exp. Psychol.* 53, 4–15. doi: 10.1027/1618-3169.53.1.4
- Shaki, S., and Fischer, M. H. (2012). Multiple spatial mappings in numerical cognition. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 804–809. doi: 10.1037/a0027562

- Shaki, S., and Fischer, M. H. (2014). Random walks on the mental number line. *Exp. Brain Res.* 232, 43–49. doi: 10.1007/s00221-013-3718-7
- Tseng, P., Bridgeman, B., and Juan, C. H. (2012). Take the matter into your own hands: a brief review of the effect of nearby-hands on visual processing. *Vis. Res.* 72, 74–77. doi: 10.1016/j.visres.2012.09.005
- Viarouge, A., Hubbard, E. M., and Dehaene, S. (2014). The organization of spatial reference frames involved in the SNARC effect. *Q. J. Exp. Psychol.* 67, 1484–1499. doi: 10.1080/17470218.2014.897358
- Vierck, E., and Kiesel, A. (2010). Congruency effects between number magnitude and response force. *J. Exp. Psychol. Learn. Mem. Cogn.* 36, 204–209. doi: 10.1037/a0018105
- Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends Cogn. Sci.* 7, 483–488. doi: 10.1016/j.tics.2003.09.002
- Weis, T., Estner, B., van Leeuwen, C., and Lachmann, T. (2016). SNARC (spatial-numerical association of response codes) meets SPARC (spatial-pitch association of response codes): automaticity and interdependency in compatibility effects. *Q. J. Exp. Psychol.* 69, 1366–1383. doi: 10.1080/17470218.2015.1082142
- Wiemers, M., Bekkering, H., and Lindemann, O. (2017). Is more always up? Evidence for a preference of hand-based associations over vertical number mappings. *J. Cogn. Psychol.* 29, 642–652. doi: 10.1080/20445911.2017.1302451
- Winter, B., Matlock, T., Shaki, S., and Fischer, M. H. (2015). Mental number space in three dimensions. *Neurosci. Biobehav. Rev.* 57, 209–219. doi: 10.1016/j.neubiorev.2015.09.005
- Wood, G., Willmes, K., Nuerk, H.-C., and Fischer, R. (2008). On the cognitive link between space and number: a meta-analysis of the SNARC effect. *Psychol. Sci. Q.* 50, 489–525.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Lohmann, Schroeder, Nuerk, Plewnia and Butz. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Spatial-Numerical Associations Enhance the Short-Term Memorization of Digit Locations

Catherine Thevenot*, Jasinta Dewi, Pamela B. Lavenex and Jeanne Bagnoud

Institute of Psychology, University of Lausanne, Lausanne, Switzerland

OPEN ACCESS

Edited by:

Hans-Christoph Nuerk,
Universität Tübingen, Germany

Reviewed by:

Koleen McCrink,
Columbia University, United States

Maria Dolores de Hevia,
Centre National de la Recherche
Scientifique (CNRS), France

Jean-Philippe van Dijck,
Ghent University, Belgium

*Correspondence:

Catherine Thevenot
catherine.thevenot@unil.ch

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 30 November 2017

Accepted: 16 April 2018

Published: 07 May 2018

Citation:

Thevenot C, Dewi J, Lavenex PB
and Bagnoud J (2018)
Spatial-Numerical Associations
Enhance the Short-Term
Memorization of Digit Locations.
Front. Psychol. 9:636.
doi: 10.3389/fpsyg.2018.00636

Little is known about how spatial-numerical associations (SNAs) affect the way individuals process their environment, especially in terms of learning and memory. In this study, we investigated the potential effects of SNAs in a digit memory task in order to determine whether spatially organized mental representations of numbers can influence the short-term encoding of digits positioned on an external display. To this aim, we designed a memory game in which participants had to match pairs of identical digits in a 9×2 matrix of cards. The nine cards of the first row had to be turned face up and then face down, one by one, to reveal a digit from 1 to 9. When a card was turned face up in the second row, the position of the matching digit in the first row had to be recalled. Our results showed that performance was better when small numbers were placed on the left side of the row and large numbers on the right side (i.e., congruent) as compared to the inverse (i.e., incongruent) or a random configuration. Our findings suggests that SNAs can enhance the memorization of digit positions and therefore that spatial mental representations of numbers can play an important role on the way humans process and encode the information around them. To our knowledge, this study is the first that reaches this conclusion in a context where digits did not have to be processed as numerical values.

Keywords: numerical cognition, space, mental number line, SNARC, short-term memory, long-term memory

INTRODUCTION

Spatial-numerical associations (SNAs) have been extensively studied since Dehaene et al.'s (1993) discovery that for individuals from Western cultures, decisions on small numbers are taken quicker with the left than the right hand, and quicker on large numbers with the right than the left hand. This SNA of response codes, or SNARC effect, was originally interpreted as the result of spatial congruency between the response hand and the position of numbers on a left-to-right mental number line (MNL) stored in long-term memory and representing increasing magnitudes of numbers (Moyer and Landauer, 1967).

The orientation of the MNL could be derived from cultural factors and especially from the direction of reading. This interpretation is supported by several studies showing that SNARC effects observed in Western participants can be reduced or even inversed in right-to-left readers (Dehaene et al., 1993; Zebian, 2005; Shaki et al., 2009). However, it has been shown that western children already preferentially represent numbers from left to right rather than right to left before school entry (Opfer et al., 2010; Thevenot et al., 2018).

For some authors, this kind of oriented representations of numbers in preliterate children stem nonetheless from environmental reading conventions (Shaki et al., 2012; Göbel et al., 2018) whereas, for others, the direction in which preschoolers count objects is not necessarily linked to their knowledge about cultural reading practice (Patro and Haman, 2017). In fact, some researchers even adopt a nativist view and argue that babies and animals share the intuition that small quantities are represented on the left of a mental-spatial continuum. For example, it has been shown that 7-month-old children prefer left-to-right arrangements of non-symbolic numerosities (sets of dots) in ascending rather than descending order (de Hevia et al., 2014) or that chicks prefer a panel placed on their left when it represents a smaller numerosity than a target, and a panel placed on their right when it represents a larger numerosity than the target (Rugani et al., 2015).

In human adults, the existence of such SNAs has been revealed by numerous studies using different paradigms (Fischer, 2001; Shaki et al., 2012; Masson et al., 2013; Masson and Pesenti, 2014; Mathieu et al., 2016, 2017; see Fischer and Shaki, 2014 for a review). For example, Fischer et al. (2003) showed that a target presented on the left side of a computer screen is detected faster when it is preceded by small rather than large numbers and, conversely, that a target on the right side of the screen is detected faster when it is preceded by larger numbers. Small and large numbers seem therefore to draw the attention of individuals to their left and right visual fields, respectively. These results have been replicated several times, notably by Bonato et al. (2009) or Dodd et al. (2008). However, Galfano et al. (2006) suggest that these attentional biases do not occur automatically but only when participants have to explicitly process the digits for numerical purposes.

In sum, even if the innate nature of the mental number line or its automaticity remains under debate, the fact that numbers and space are associated is now well established. However, as noted by McCrink and Shaki (2016), less is known about how SNAs affect the way individuals process their environment, especially in terms of learning and memory. To address this question, we have investigated the potential effects of SNAs in a memory task. Specifically, we aimed to determine whether the spatial position of digits is better memorized when they are congruent rather than incongruent with the positions of numbers on the internalized MNL.

This question was recently addressed by Gut and Staniszewski (2016) who presented digits to the right or to the left of a central fixation point. When the digits disappeared, the fixation point was replaced by one of the two digits, and participants had to recall the location where the target digit had been previously presented. The authors showed that memory performance (i.e., shorter RTs and lower error rates) was better for “small” as compared to “large” digits when they were positioned to the left of the fixation point. However, the reverse congruency effect was not observed. In other words, when the digits were presented to the right of the fixation point, both “small” and “large” digits led to similar memory performance. The authors offered several explanations for these results. First, they suggested that a congruency effect only for small numbers constitutes evidence

that the magnitude of numbers plays a role in SNARC-like effects. According to the authors, small numbers such as one or two might be encountered more often and might better catch individuals’ attention than larger numbers (Cai and Li, 2015) and this would partly explain why congruency effects were only obtained for very small one-digit numbers. Nevertheless, the authors note that, in opposition to these assumptions, smaller numbers might not be encountered more often than larger numbers (e.g., Dehaene and Mehler, 1992) but that individuals might react faster to higher than lower magnitude numbers (e.g., Krause et al., 2017). Moreover, Gut and Staniszewski (2016) observed that, independent of number magnitudes, their participants were faster to recall the location of digits when they were displayed to the right rather than to the left of the fixation point. However, because digits to the right of the fixation point were always responded to with the right hand (and digits to the left of the fixation point responded to with the left hand) and given the fact that all their participants were right-handed, better dexterity with the right hand might explain their results. All in all, and despite the interesting question raised by the authors, a coherent explanation of their findings was not obvious. Their conclusion that “the spatial representation of numbers on the MNL are crucial for retrieval of numbers presented on the left and that the responses to the numbers presented on the right are generally faster and more correct irrespective of their congruency” (p. 203) is difficult to reconcile with the results of the numerous studies showing faster processing for large numbers when they are presented in the right visual field of participants, including the famous original SNARC effect.

The question of potential effects of digit processing in a memory task has also been addressed by McCrink and Galamba (2015) in a series of experiments in which spatial locations had to be memorized. Participants were presented with a series of sequentially highlighted spatial locations on a grid and their task was to repeat the sequence by touching the locations on a computer screen. The locations could appear from left to right, from right to left or randomly in the grid. In one of the conditions, symbolic numerals were associated with the locations but, contrary to the authors’ expectations, there was no advantage of the left-to-right over the right-to-left flow for the recall of the locations. However, it is possible that in McCrink and Galamba’s (2015) task, the sequential movement of the digits in a two-dimensional space might preclude them from being influenced by the MNL where number are represented strictly one-dimensionally.

In sum, the question of whether there are improvements in the encoding and recall of numbers when their spatial positions are congruent with MNL orientation remains unanswered. The object of the current study is to further investigate this question with a memory game in which adult participants had to match pairs of identical numbers. The game was presented on a computer screen where two rows of cards hiding numbers were displayed. The first row of nine cards was created by using the nine digits from 1 to 9. In the congruent condition, the digits 1–4 were randomly placed to the left of the five, which was placed in the exact middle of the row, and the digits 6–9 were randomly placed to its right. In the incongruent condition, small

digits from 1 to 4 were placed to the right of the five and larger digits from 6 to 9 were placed to its left. Finally, in the random condition, digits were randomly placed on the first row. For all the configurations, the digits from 1 to 9 were pseudo-randomly positioned on the second row. If the organization of numbers on the MNL can enhance the memorization of numerical material, individuals should perform better in the congruent than in the incongruent and random conditions. Moreover, and conversely, if the organization of numbers on the MNL can interfere with the memorization of numerical material, individuals should also perform better in the random than in the incongruent condition in which the presentation of numbers conflicts maximally with the representations of numbers on the MNL.

MATERIALS AND METHODS

Participants

Twenty-five right-handed undergraduate students in Psychology at the University of Lausanne participated for course credit. Participants were aged between 18 and 32 years (mean: 21.64 years) and five of them were men.

Material and Procedure

The task we designed was an adaptation of the classic memory game where cards are placed face down in front of a player who has to find matching pairs. To do so, he or she has to turn the cards face up and encode the symbol that occurs on the card and its position before turning them face down again. In our adaptation, participants were presented with two rows of nine squares, representing the cards, on a computer screen. The symbols that had to be encoded were digits from 1 to 9. Participants were first introduced to the task and instructed how to play with a shorter version of the task using geometrical shapes instead of digits.

During each game, participants were presented with two rows of cards that they could turn face up and down by clicking on them. The participants had been instructed that the first card to turn face up was the card at the leftmost position of the first row. Once the digit was seen, the card had to be turned face down again and the card immediately to its right had to be turned face up. This rule had to be applied until the end of the game and, at the end of the first row, it was the card at the leftmost position of the second row that had to be turned over. However, if the participant thought that she had already seen the digit on the card, she could try to find it by returning to and overturning a previous card. If she succeeded, the two cards stayed visible on the screen and she could continue with the game. If she made a mistake, she had to turn the last card (the incorrect choice) face down and then could either attempt to find the matching card again, or she could give up for this pair and continue the game by turning over the next card in the row. Participants were not aware that it was not possible to find a pair before the end of the first row. If all the pairs were not found when the last card was turned up, the participant was free to return any card of the game. A perfect game, without any mistakes, could be completed in 27 moves, corresponding to nine moves on the first line to discover the positions of the nine

digits and 18 moves to match the pairs (i.e., one move to overturn each card on the second row and one move to match it to a card in the first row).

As described in the Introduction, the first row of cards could be in a congruent, incongruent or random arrangement. In the congruent condition, the digits 1–4 were randomly placed to the left of the five, which was placed in the middle of the row, and the digits 6–9 were randomly placed to its right. In the incongruent condition, the digits of the congruent condition were replaced by digits using the inverse ordering of numbers (i.e., 1 replaced by 9, 2 replaced by 8, 3 replaced by 7, 4 replaced by 6 and vice versa). Finally, in the random condition, digits were all placed pseudo-randomly in the first row with the rule that all small or all large digits could not be positioned on the same side of the row. For all the conditions, the digits from 1 to 9 were pseudo-randomly positioned on the second row with the rule that two matching symbols were separated at least by five sequential positions (e.g., the number six could not be the last digit on row one and the first digit on row two). In order to minimize the risk of accidental biases, two versions of the randomized material were created for each condition and all participants played the same randomized versions of the game in all three conditions (**Figure 1**). Therefore, participants played the game six times (3 conditions \times 2 versions). For each of the versions, the three conditions were presented in a counterbalanced manner to participants, so that 1/3 played the game with the congruent condition first, 1/3 played the game with the incongruent condition first, and 1/3 played the game with the random condition first. For each participant, we measured the time required to complete each of the six games and the number of moves realized to do so.

RESULTS

Solution Times and Accuracy

A repeated measures ANOVA with configuration (congruent, incongruent, and random) as a within-subjects factor revealed differences in the solution times, $F(2,48) = 4.81$, $\eta_p^2 = 0.17$, $p = 0.01$, and the number of moves, $F(2,48) = 4.44$, $\eta_p^2 = 0.16$, $p = 0.02$ (**Table 1**).

Because the SNA hypothesis allowed us to put forward precise predictions, one-sided planned comparisons with Bonferroni corrections were conducted to compare solution times between the congruent condition and the other two conditions. They revealed that solution times were shorter in the congruent (55.71 s) than in the incongruent (64.95 s), $z = 2.62$, $p = 0.01$, and random conditions (65.41 s), $z = 2.75$, $p < 0.01$. For the comparison between the incongruent and random conditions, a one-sided planned comparison with Bonferroni correction showed no significant difference between the two conditions, $z = 0.13$, $p = 1$. An additional Bayesian analysis on this difference revealed substantial evidence for this absence of effect ($BF_{10} = 0.28$).

The same pattern of results was obtained for the number of moves. Indeed, participants completed the game in a fewer number of moves in the congruent (31.86 moves) than in the

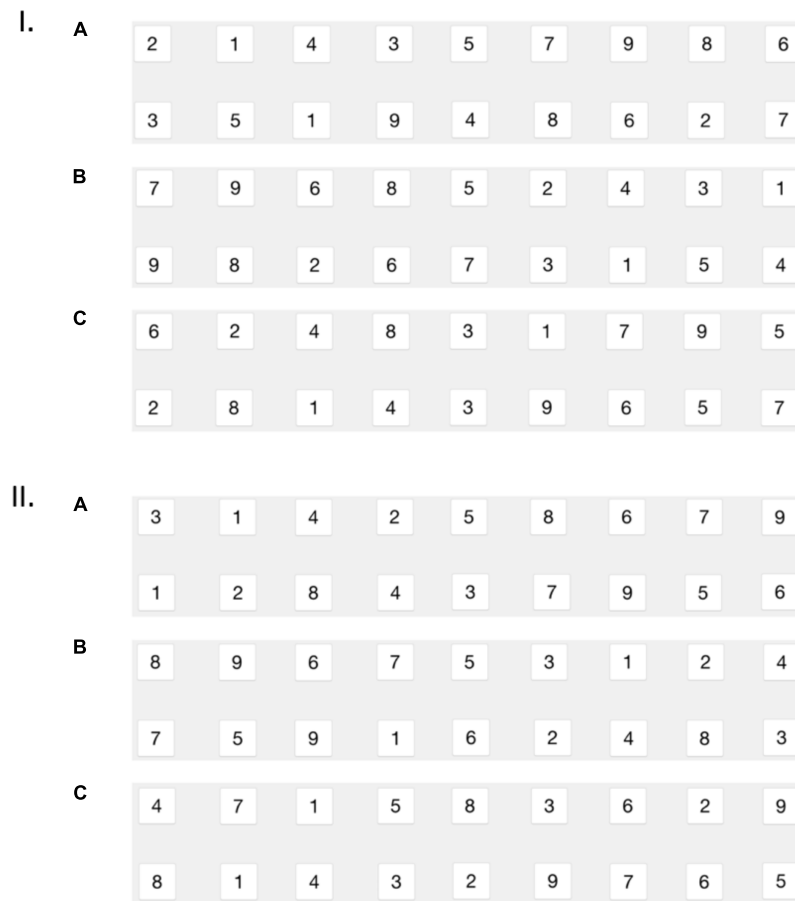


FIGURE 1 | The two versions of the randomized material in the memory game (I and II) with the three different configurations: congruent (A), incongruent (B), and random (C).

TABLE 1 | Solution times (in seconds) and number of moves (and standard-deviations) for the congruent, incongruent, and random configurations in the memory game.

	Solution times	Number of moves
Congruent	55.71 (14.74)	31.86 (3.59)
Incongruent	64.95 (24.43)	34.64 (7.24)
Random	65.41 (18.49)	35.62 (5.39)

incongruent (34.64 moves), $z = 2.12$, $p = 0.05$, and random conditions (35.62 moves), $z = 2.87$, $p < 0.01$. Again, there was no significant difference in the number of moves between the incongruent and random conditions, $z = 0.75$, $p = 0.68$ and the Bayesian analysis on this difference revealed substantial evidence for this absence of effect ($BF_{10} = 0.325$).

DISCUSSION

In this research we were interested in the question of whether representations of numbers on a mental number line can influence the short-term memorization of number positions

displayed in front of a participant. To this aim, we adapted the classic memory game and showed that the positions of digits representing smaller and larger numbers were more easily recalled when they were presented on the left and on the right of the display, respectively, than when it was the reverse or when the digits were randomly positioned. This finding suggests that the spatial mental representation of numbers in a left-to-right orientation can facilitate the memorization of digit localizations. An oriented representation of numbers activated during the memory task could have indeed constituted a framework that helped individuals to encode and recall the positions of the digits. Importantly, in the task that we designed the digits to be memorized did not need to be processed as numbers. It seems therefore that SNAs can be activated in absence of explicit processing of number magnitudes. Interestingly, these congruency effects were observed despite possible interferences due to “micro-incongruences” on both sides of our display. In fact, due to the constraints of the task, the digits could not be presented strictly in the canonical order and therefore, within one side, a smaller digit could be preceded or followed by a larger digit. Obviously these “micro-incongruences,” which were present in the three

conditions of our experiment, did not significantly impact our results.

Contrary to our expectations, we did not observe differences in memory performance between the random and the incongruent conditions. Thus, it appears that a presentation of numbers that maximally conflict with the organization of the MNL is not more detrimental to memory performance than a less incongruent (e.g., random) presentation of numbers. This result may lead to a number of different interpretations. First, it is possible that the MNL is not automatically activated as soon as numerals are presented to the cognitive system and that it is activated only when the presented configurations match an ordered sequence of numbers stored in long-term memory. In other words, the MNL would not be activated when numbers are encountered in an order that does not match any long-term memory representations, hence the lack of difference between the incongruent and the random conditions. According to this interpretation, it would not even be necessary to assume that numbers are organized from left-to-right in long-term memory but simply that numbers are ordered in long-term memory. Indeed, according to an alternative account of SNARC effects, associations between space and numbers are due to the characteristics of maintenance of information in short-term memory rather than magnitude representations in long-term memory (e.g., van Dijck and Fias, 2011; van Dijck et al., 2014; Abrahamse et al., 2016). A strong empirical argument for this view is that any ordered information maintained in short-term memory, such as months of the year, letters of the alphabet or even a list of words just memorized, is represented from left to right and is subject to SNARC effects (e.g., Gevers et al., 2003, 2004). Within this framework, the organization of numbers in the congruent condition of our experiment would convoke a non-oriented ordered sequence of numbers from long-term memory and the ordered sequence would temporarily be oriented from left to right in short-term memory. This transitory representation of numbers would serve as a framework to encode and recall the position of digits. In this case, our results would contradict Dehaene's seminal interpretation of SNARC effects (Dehaene et al., 1993) according to which the magnitude of number is automatically activated and inherently associated with space.

An alternative interpretation of the absence of differences in memory performance between the incongruent and the random conditions could be that in both cases the MNL is automatically activated but that the difficulty to memorize random configurations equates with the difficulty to inhibit information in total conflict with the MNL organization. Further experiments designed to directly contrast these alternative interpretations will have to be conducted in the future. One possibility would be to examine potential MNL priming effects across conditions. Indeed, when the incongruent and random conditions are presented before the congruent condition, and if the MNL is automatically activated in these conditions, priming effects of the MNL should be observed in the subsequent congruent condition. The improvement in memory performance that we observed in the congruent condition of our experiment should therefore be increased. Conversely, even

if the MNL is not automatically activated in the incongruent and random conditions, it is likely to be activated after participants performed the task in the congruent condition. In this case, we should observe priming effects of the MNL on the incongruent and random conditions when they are performed after the congruent condition but no priming effects of the MNL on the congruent condition when it is performed before the random and incongruent conditions. Unfortunately, such analyses are impossible with the present data set because participants played too few games in each condition.

Finally, a last alternative interpretation of our results is that enhanced memory performance in the congruent condition is not due to any ordering of numerical information in short-term memory but only to the number themselves, which could trigger the attention of participants on the left or of the right attentional fields depending on their size (i.e., on the left for small numbers and on the right for larger ones) (Gevers et al., 2006; Proctor and Cho, 2006; Santens and Gevers, 2008). Nevertheless, we think that the lack of difference in memory performance between the incongruent and random conditions argues against this interpretation. Indeed, the conflicts between attentional biases triggered by the magnitude of numbers and the position of the numbers in the memory game are maximal in the incongruent condition and memory performance should therefore be worst in this condition than in the random condition. Still, this line of reasoning is based on a lack of effect and has to be considered with care.

CONCLUSION

We have shown that SNAs can enhance the memorization of digit positions and thus that spatial mental representations of numbers could play an important role in the way that humans process and encode the information around them. To our knowledge, this study is the first that reaches this conclusion in a context of a memory task where the digits did not have to be processed as numerical values. This suggests that either Arabic numerals cannot be perceived as pure symbols lacking numerical characteristics, or that individuals can consciously evoke numerical knowledge associated with digits when it is potentially helpful for them.

ETHICS STATEMENT

Signed informed consent forms were obtained from all our participants before they entered the study. The experiment reported in the present manuscript has been conducted in compliance with the Swiss Law on Research involving human beings and because only behavioral data were collected in a non-vulnerable population of adults, the approval of the Canton de Vaud ethic committee was not required. This study was carried out in accordance with the recommendations of the Ethics Committee of the University of

Lausanne with written informed consent from all subjects in accordance with the Declaration of Helsinki.

JB conducted the experiment. All the authors contributed to the conception of the experiment.

AUTHOR CONTRIBUTIONS

CT had the original idea of the research. CT and JB wrote a first draft of the manuscript and PL edited the manuscript.

FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

- Abrahamse, E., van Dijck, J.-P., and Fias, W. (2016). How does working memory enable number-induced spatial biases? *Front. Psychol.* 7:977. doi: 10.3389/fpsyg.2016.00977
- Bonato, M., Priftis, K., Marenzi, R., and Zorzi, M. (2009). Normal and impaired reflexive orienting of attention after central nonpredictive cues. *J. Cogn. Neurosci.* 21, 745–759. doi: 10.1162/jocn.2009.21054
- Cai, Y. C., and Li, S. X. (2015). Small number preference in guiding attention. *Exp. Brain Res.* 233, 539–550. doi: 10.1007/s00221-014-4134-3
- de Hevia, M. D., Girelli, L., Addabbo, M., and Macchi Cassia, V. (2014). Human infants' preference for left-to-right oriented increasing numerical sequences. *PLoS One* 9:e96412. doi: 10.1371/journal.pone.0096412
- Dehaene, S., Bossini, S., and Giraux, P. (1993). The mental representation of parity and number magnitude. *J. Exp. Psychol.* 122, 371–396. doi: 10.1037/0096-3445.122.3.371
- Dehaene, S., and Mehler, J. (1992). Cross-linguistic regularities in the frequency of number words. *Cognition* 43, 1–29. doi: 10.1016/0010-0277(92)90030-L
- Dodd, M. D., Van der Stigchel, S., Leghari, M. A., Fung, G., and Kingstone, A. (2008). Attentional SNARC: There's something special about numbers (let us count the ways). *Cognition* 108, 810–818. doi: 10.1016/j.cognition.2008.04.006
- Fischer, M. H. (2001). Number processing induces spatial performance biases. *Neurology* 57, 822–826. doi: 10.1212/WNL.57.5.822
- Fischer, M. H., Castel, A. D., Dodd, M. D., and Pratt, J. (2003). Perceiving numbers causes spatial shifts of attention. *Nat. Neurosci.* 6, 555–556. doi: 10.1038/nn1066
- Fischer, M. H., and Shaki, S. (2014). Spatial associations in numerical cognition - From single digits to arithmetic. *Q. J. Exp. Psychol.* 67, 1461–1483. doi: 10.1080/17470218.2014.927515
- Galfano, G., Rusconi, E., and Umiltà, C. (2006). Number magnitude orients attention, but not against one's will. *Psychon. Bull. Rev.* 13, 869–874. doi: 10.3758/BF03194011
- Gevers, W., Reynvoet, B., and Fias, W. (2003). The mental representation of ordinal sequences is spatially organised. *Cognition* 87, B87–B95. doi: 10.1016/S0010-0277(02)00234-2
- Gevers, W., Reynvoet, B., and Fias, W. (2004). The mental representation of ordinal sequences is spatially organised: evidence from days of the week. *Cortex* 40, 171–172. doi: 10.1016/S0010-9452(08)70938-9
- Gevers, W., Verguts, T., Reynvoet, B., Caessens, B., and Fias, W. (2006). Numbers and space: a computational model of the SNARC effect. *J. Exp. Psychol.* 32, 1197–1207. doi: 10.1037/0096-1523.32.1.32
- Göbel, S. M., McCrink, K., Fischer, M. H., and Shaki, S. (2018). Observation of directional storybook reading influences young children's counting direction. *J. Exp. Child Psychol.* 166, 49–66. doi: 10.1016/j.jecp.2017.08.001
- Gut, M., and Stanislawski, R. (2016). The SNARC effect in number memorization and retrieval. What is the impact of congruency, magnitude and the exact position of numbers in short-term memory processing? *Adv. Cogn. Psychol.* 12, 193–208. doi: 10.5709/acp-0198-0
- Krause, F., Bekkering, H., Pratt, J., and Lindemann, O. (2017). Interaction between numbers and size during visual search. *Psychol. Res.* 81, 664–677. doi: 10.1007/s00426-016-0771-4
- Masson, N., and Pesenti, M. (2014). Attentional bias induced by solving simple and complex addition and subtraction problems. *Q. J. Exp. Psychol.* 67, 1514–1526. doi: 10.1080/17470218.2014.903985
- Masson, N., Pesenti, M., and Dormal, V. (2013). Spatial bias in symbolic and non-symbolic numerical comparison in neglect. *Neuropsychologia* 51, 1925–1932. doi: 10.1016/j.neuropsychologia.2013.06.004
- Mathieu, R., Epinat-Duclos, J., Sigovan, M., Breton, A., Cheylus, A., Fayol, M., et al. (2017). What's behind a "+" sign? Perceiving an arithmetic operator recruits brain circuits for spatial orienting. *Cereb. Cortex* doi: 10.1093/cercor/bhx064 [Epub ahead of print].
- Mathieu, R., Gourjon, A., Couderc, A., Thevenot, C., and Prado, J. (2016). Running the number line: operators elicit horizontal shifts of attention during single-digit arithmetic. *Cognition* 146, 229–239. doi: 10.1016/j.cognition.2015.10.002
- McCrink, K., and Galamba, J. (2015). The impact of symbolic and non-symbolic quantity on spatial learning. *PLoS One* 10:e0119395. doi: 10.1371/journal.pone.0119395
- McCrink, K., and Shaki, S. (2016). Culturally inconsistent spatial structure reduces learning. *Acta Psychol.* 169, 20–26. doi: 10.1016/j.actpsy.2016.05.007
- Moyer, R. S., and Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature* 215, 1519–1520. doi: 10.1038/2151519a0
- Opfer, J. E., Thompson, C. A., and Furlong, E. E. (2010). Early development of spatial-numeric associations: evidence from spatial and quantitative performance of preschoolers. *Dev. Sci.* 13, 761–771. doi: 10.1111/j.1467-7687.2009.00934.x
- Patro, K., and Haman, M. (2017). Playing reading, using hands: Which activities are linked to number-space processing in preschool children? *Psychol. Lang. Commun.* 21, 325–351. doi: 10.1515/plc-2017-0016
- Proctor, R. W., and Cho, Y. S. (2006). Polarity correspondence: a general principle for performance of speeded binary classification tasks. *Psychol. Bull.* 132, 416–442. doi: 10.1037/0033-2909.132.3.416
- Rugani, R., Vallortigara, G., Priftis, K., and Regolin, L. (2015). Number-space mapping in the newborn chick resembles humans' mental number line. *Science* 347, 534–536. doi: 10.1126/science.aaa1379
- Santens, S., and Gevers, W. (2008). The SNARC effect does not imply a mental number line. *Cognition* 108, 263–270. doi: 10.1016/j.cognition.2008.01.002
- Shaki, S., Fischer, M. H., and Göbel, S. M. (2012). Direction counts: a comparative study of spatially directional counting biases in cultures with different reading directions. *J. Exp. Child Psychol.* 112, 275–281. doi: 10.1016/j.jecp.2011.12.005
- Shaki, S., Fischer, M. H., and Petrusic, W. M. (2009). Reading habits for both words and numbers contribute to the SNARC effect. *Psychon. Bull. Rev.* 16, 328–331. doi: 10.3758/PBR.16.2.328
- Thevenot, C., Fayol, M., and Barrouillet, P. (2018). Spatial numerical associations in preschoolers. *Think. Reason.* 24, 221–233. doi: 10.1080/13546783.2017.1375013
- van Dijck, J. P., Abrahamse, E. L., Acar, F., Ketels, B., and Fias, W. (2014). A working memory account of the interaction between numbers and spatial attention. *Q. J. Exp. Psychol.* 67, 1500–1513. doi: 10.1080/17470218.2014.903984
- van Dijck, J. P., and Fias, W. (2011). A working memory account for spatial-numerical associations. *Cognition* 119, 114–119. doi: 10.1016/j.cognition.2010.12.013
- Zebian, S. (2005). Linkages between number concepts, spatial thinking, and directionality of writing: the SNARC effect and the REVERSE SNARC effect in English and Arabic monoliterates, biliterates, and illiterate Arabic speakers. *J. Cogn. Cult.* 5, 165–190. doi: 10.1163/1568537054068660

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Thevenot, Dewi, Lavenex and Bagnoud. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Spatial and Verbal Routes to Number Comparison in Young Children

Francesco Sella^{1*}, Daniela Lucangeli² and Marco Zorzi³

¹ Department of Experimental Psychology, University of Oxford, Oxford, United Kingdom, ² Department of Developmental Psychology and Socialisation, Università di Padova, Padova, Italy, ³ Department of General Psychology, Università di Padova, Padova, Italy

The ability to compare the numerical magnitude of symbolic numbers represents a milestone in the development of numerical skills. However, it remains unclear how basic numerical abilities contribute to the understanding of symbolic magnitude and whether the impact of these abilities may vary when symbolic numbers are presented as number words (e.g., “six vs. eight”) vs. Arabic numbers (e.g., 6 vs. 8). In the present study on preschool children, we show that comparison of number words is related to cardinality knowledge whereas the comparison of Arabic digits is related to both cardinality knowledge and the ability to spatially map numbers. We conclude that comparison of symbolic numbers in preschool children relies on multiple numerical skills and representations, which can be differentially weighted depending on the presentation format. In particular, the spatial arrangement of digits on the number line seems to scaffold the development of a “spatial route” to understanding the exact magnitude of numerals.

OPEN ACCESS

Edited by:

Maciej Haman,
University of Warsaw, Poland

Reviewed by:

Melissa M. Kibbe,
Boston University, United States
Jennifer B. Wagner,
College of Staten Island,
United States

*Correspondence:

Francesco Sella
sella.francesco@gmail.com

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 05 October 2017

Accepted: 01 May 2018

Published: 24 May 2018

Citation:

Sella F, Lucangeli D and Zorzi M
(2018) Spatial and Verbal Routes
to Number Comparison in Young
Children. *Front. Psychol.* 9:776.
doi: 10.3389/fpsyg.2018.00776

Keywords: counting, numerical estimation, number line task, digit comparison, preschool children

INTRODUCTION

A wealth of studies have established an intimate association between numbers and space (Hubbard et al., 2005; de Hevia et al., 2008; Nuerk et al., 2015; Patro et al., 2016). This association emerges early in development, as attested by the finding that 7 months-old infants display preferential looking for increasing numerical magnitude from left-to-right (De Hevia et al., 2014). Preschool children also associate small numerosities with the left side of space and large numerosities with the right side of space (Patro and Haman, 2012; see also, Patro et al., 2016). Interestingly, a spontaneous association between numerical quantity and space has also been found in new-born chicks (Rugani et al., 2015) as a sign of an evolutionarily ancient link.

Symbolic numbers are also strongly related to space as shown by the association between relatively small numbers with the left side of space and relatively large numbers with the right side of space (the SNARC effect; Dehaene et al., 1993). Patients with spatial neglect, who fail to pay attention to the left side of the visual field, also neglect small numbers when asked to verbally bisect numerical intervals (Zorzi et al., 2002). This has suggested that numerical magnitudes are mentally represented in a spatially ordered manner along a putative Mental Number Line (Restle, 1970; Dehaene et al., 1993; Zorzi et al., 2002) and that number processing involves orienting of attention in this “number space” (Zorzi et al., 2002, 2012; Fischer et al., 2003; Hubbard et al., 2005; Umiltà et al., 2009).

The ability to map symbolic numbers onto spatial positions has been extensively studied in developmental studies on primary school children using Siegler and Opfer’s (2003) number-to-position task (Siegler and Booth, 2004; Booth and Siegler, 2006; Siegler et al., 2009). More recently, Sella et al. (2017) observed that spatial mapping of symbolic numbers emerges during the

early preschool period (also see, Berteletti et al., 2010) and appears to be crucial for understanding magnitude relationships for exact numbers. The aim of the present study was to further investigate how the ability to map numbers on a visual horizontal line is linked to symbolic number comparison skills.

Note that the understanding of symbolic numbers is typically linked to the development of counting. Around the age of two, toddlers begin to implement the counting routine to enumerate objects in their environment (Wynn, 1992). Children have to respect three foundational principles to achieve correct counting (Gelman and Gallistel, 1978): reciting the number words sequence in the established order (*stable order* principle); matching each object in the set to one and only one number word (*one-to-one correspondence* principle); identifying that the last number word represents the numerosity of the set (*cardinality* principle). Acquisition of counting principles is a long and error-prone process that engages children for about 1½ years, usually between 2 and 4 years of age (Sarnecka, 2015). According to the knower-level theory (Wynn, 1990; Carey, 2001; Sarnecka and Carey, 2008), children initially lack the understanding of number words: When requested to collect a certain number of objects (as in the Give-a-number task; Wynn, 1990), these children usually “grab” a handful of items without implementing any structured counting procedure. Subsequently, children sequentially learn the cardinal meaning of the number words from “one” to “four” and are able to provide numerosities from one to four when requested. These children are usually defined as Subset-knowers because their cardinal meaning of number words is limited to a subset of the counting list. Finally, children understand that the next number word in the counting list corresponds to one additional element in the counted set (i.e., $n + 1$, Gelman and Gallistel, 1978). Children at this stage can extend the cardinality principle to the entire counting list, thereby becoming Cardinal-Principle knowers (CP-knowers).

The acquisition of the cardinality principle should allow children to correctly map number words to corresponding objective external numerosities and, therefore, to understand the magnitude relation between number words (i.e., “eight is more than six”). Nevertheless, the acquisition of the cardinality principle does not imply a full understanding of the magnitude relation between number words. Indeed, some CP-knowers can fail in choosing the larger number when confronted with a pair of number words with magnitudes greater than 4 (e.g., 8 vs. 6), although they are successful when at least one number in the pair belongs to the small number range (≤ 4) (e.g., 4 vs. 2 or 6 vs. 3) (Le Corre, 2014). The paradox emerges from the fact that CP-knowers can reliably count both small and large numerical sets, as in the Give-a-number task, thereby showing the ability to connect number words to the corresponding external numerical quantities. Le Corre (2014) observed that the ability to compare pairs of large number words was present only in a subset of CP-knowers who were also able to reliably estimate large (i.e., >4 items) briefly visually presented numerosities (i.e., numerosity estimation). These children were referred to as CP-mappers, because their ability to map external numerosities onto number words is not derived by merely implementing the counting routine. Accordingly, these children know that later

number words in the counting list are associated with larger numerical quantities (i.e., later-greater principle) and, then use this knowledge to determine the larger between two number words.

Sella et al. (2017) used a similar approach to investigate the relation between the acquisition of the cardinality principle and spatial mapping of numbers in a sample of preschool children. CP-knowers were classified as mappers when they could reliably place numerals on the horizontal visual line in the number-to-position task (1–10 interval) and as non-mappers when their positioning lacked any numerical meaning (e.g., all the numbers placed in the middle of the line). Crucially, only CP-mappers proficiently chose the larger between two visually presented Arabic digits whereas CP-non-mappers’ performance was close to chance level. Note that the spatial arrangement of digits on the number line is a powerful source of information because the magnitude of a digit can be conveyed by its location in relation to the location of other digits. Children who have internalized the spatial disposition of digits and understood that spatial shifts along the line represent changes in magnitude (*spatial mapping principle*; Sella et al., 2017) can use this information to infer the magnitude of numerals and compare them.

In summary, the magnitude comparison of number words seems to relate to the ability to map external numerical quantities onto the counting list (Le Corre, 2014), whereas the ability to compare visually presented digits may be linked to the ability to spatially map numbers (Sella et al., 2017). In the case of number words, the ability to linearly map external numerosities to the counting list marks the understanding that the later number words in the counting list are associated with larger numerical quantities. For Arabic digits, instead, the ability to map them to space informs about the magnitude of digits based on their absolute position on the line and their relative position compared to other digits. However, it remains unclear whether the contribution of numerosity estimation and spatial mapping are tied to a specific presentation format or are both related to the understanding of the magnitude relation between symbolic numbers.

More broadly, the investigation of format-dependent acquisition of the numerical meaning of symbols in young children is rather sparse. Some authors have suggested that children independently associate number words and Arabic digits to the corresponding numerical quantities and later number words are mapped to Arabic digits (Benoit et al., 2013). Others, instead, have suggested that the mapping between number words and Arabic digits is learnt after the mapping between number words and numerical quantities (Hurst et al., 2016). Interestingly, CP-knowers fail in transferring the cardinality knowledge of number words to Arabic digits, even though they can correctly read Arabic digits, thereby converting them from the visual to the verbal format (Knudsen et al., 2015). A recent detailed investigation of the mapping between number words, Arabic digits, and numerical quantities highlighted that the mapping between digits and numerical quantities contributed to the digit comparison performance, with an indirect contribution of the word-digit and word-quantity mappings (Jiménez Lira et al., 2017). Overall, these results

suggest the existence of a separate (visual) route for learning the numerical meaning of Arabic digits, which coexists in parallel to the learning of the numerical meaning of number words. Nevertheless, it is still plausible that children initially learn the numerical meaning of number words and subsequently transfer this knowledge to Arabic digits while learning to read them.

The aim of the present within-subjects study was to investigate this issue in relation to children's ability to compare number words and Arabic digits. Assessing whether a core numerical skill, like symbolic number comparison, is modulated by the presentation format can inform theories of numeracy development and might have an impact on educational practices. Our hypothesis that symbolic number comparison in young children relies on distinct routes (spatial vs. verbal) depending on the presentation format leads to specific predictions. That is, performance in number words comparison should be related to the ability to estimate large numerical quantities (Le Corre, 2014) after controlling for cardinality knowledge, whereas accuracy of spatial mapping should be irrelevant. Conversely, performance in Arabic digit comparison should be related to the accuracy of spatial mapping after controlling for cardinality knowledge (Sella et al., 2017), whereas the precision in estimating large numerical quantities should be irrelevant. It is worth considering that children may transform the Arabic digit comparison into a number words comparison by transcoding the Arabic code into verbal code (Dehaene, 1992). If that is the case, the ability to read digits and performance in the number words comparison task should explain the performance in the Arabic digit comparison task and the role of the accuracy in spatial mapping should be minimal.

MATERIALS AND METHODS

Participants

Sixty preschool children from a school located in north-eastern Italy took part in the experiment after informed consent was obtained from parents or legal guardians. Seven children were removed from the analyses because they failed to complete the experimental session (three children interrupted the session and one child provided only three estimates in the numerosity estimation task) or they had a cognitive disability as reported by the teachers (three children). Six additional participants were removed from analyses because they failed to correctly recite the numerical sequence at least up to 10 in the forward enumeration task (see below), which was a crucial requirement to perform the numerosity estimation task (which contained trials with numerosity up to 10). The final sample was composed of 47 children (17 boys, $M_{\text{age-in-months}} = 64$, $SD = 9$, range = 43–79), a sample size that is in line with those of the relevant previous studies (Le Corre, 2014; Sella et al., 2017).

Procedure

Children were met individually in a separate quiet room during school hours and completed all the tasks in one experimental session (approximately 20–30 min depending on the child's ability). Children completed the numerical tasks in the following

order: forward enumeration, backward enumeration, give-a-number, naming, number line, Arabic digit comparison, number words comparison and numerosity estimation. Children were allowed to take a break between tasks and they could interrupt the experimental session at any time. The results from the backward enumeration task are not reported in the present study.

Numerical Tasks

Forward Enumeration

Children were asked to recite the numerical sequence starting from one and were stopped when they reached 50 or when they could not go any further. Children could correct themselves immediately if they realized they have committed a mistake. The experimenter did not provide any feedback. This task was administered to ensure that children were at least able to recite the counting list up to 10, which was the largest numerosity presented in the numerosity estimation task (see below).

Give-a-Number (GaN)

A small basket with 15 wooden tomatoes (approximately 3 cm of diameter) was at the child's disposal before starting the task in order to familiarize the child with the materials. The task was introduced as a role-play game in which the experimenter played the role of a customer and the child played the role of the grocer. The experimenter said: *"Let's play the market game! You are a grocer and I'm a customer that wants to buy some delicious tomatoes. Ok? Are you ready?"* The experimenter then said: *"Hello! May I have n tomato/es, please?"* As soon as the child gave the selected number of tomatoes, the experimenter said: *"Is this/Are these n tomato/es?"* The child was allowed to modify the number of tomatoes until she was sure about the number. The experimenter asked for 1, 2, 3, 4, 5, 8, and 10 tomatoes in random order and the percentage of correct responses was calculated.

Naming

Children were presented with an Arabic digit in the center of the computer screen and were asked to name it aloud. Numbers from 0 to 20 were presented randomly. Only digits from 1 to 9 were considered given that the same range of digits was presented in the Arabic digit comparison task, in which children were presented with digits that were not read by the experimenter. One point was awarded for each correct naming and the percentage of correct responses was calculated.

Number Lines 1–10 (NL)

A black horizontal line, with no tick marks, was presented in the middle of the computer screen with the number one ("1") placed just below the left-end of the line and the number ten ("10") placed just below the right-end. The number to be positioned (e.g., "4") was presented inside a box in the upper left corner of the screen. For every trial, the experimenter said: *"This line goes from one to ten [pointing at the numbers]. Where is the correct place for n [pointing at the number in the upper left corner]? Show me the correct place moving the mouse and pressing the mouse button when you are on the right place!"* Children placed the numbers on the line by moving an arrow using the mouse and clicking the mouse button to confirm the selected position. The

movement of the arrow was constrained to the horizontal line to facilitate the response. After pressing the mouse button, a red dot appeared on the selected location. There were two training trials (i.e., 1 and 10) in which, if the positioning of the target number was not accurate, the experimenter indicated to the child the correct position. The experimenter intervened only 4 times to correct children in the training trials. Out of 47 children included in the study, 44 correctly placed the number 1, one child placed 1 close to the position of 2 and two children placed 1 almost in the position of 10. Forty-six children correctly placed the number 10 and one child placed it close to the position of 9. After the training trials, children had to place eight randomly presented numbers (i.e., 2, 3, 4, 5, 6, 7, 8, and 9), three times each for a total of 24 trials. For each child, we calculated the mean percentage of absolute error (PAE) as follows: $(|\text{estimate} - \text{target number}|/9) \times 100$. We also calculated the individual regression slope of estimates as function of target numbers ($M = 0.69$, $SD = 0.45$, range: -0.40 , 1.27): children with a positive and significant regression slope were classified as *spatial mappers* ($n = 34$; $M = 0.92$, $SD = 0.25$, range: 0.25 , 1.27) whereas the remaining children were classified as *non-mappers* ($n = 13$; $M = 0.09$, $SD = 0.25$, range: -0.40 , 0.39).

Number Comparison

Number words comparison (adapted from Le Corre, 2014)

Two gray boxes were horizontally presented in the lowest part of the computer screen. Then, the experimenter read the text written above the boxes: “In this box [pointing the box on the left side] there is/are n ball/s and in this box [pointing the box on the right side] there is/are m ball/s. Which box has more balls?” The child responded by pointing the box (or simply saying which was the largest number) and the experimenter recorded the response by pressing the left or right button of the touchpad. After the response, the two boxes were replaced by two images showing the actual numerosities. The images representing the comparison numerosities were generated following a method to control for the influence of physical variables (e.g., cumulative surface area, convex hull; Gebuis and Reynvoet, 2011). Then, the experimenter read the text written above the boxes: “This box [pointing the box on the left side] contained n ball/s and this box [pointing the box on the right side] contained m ball/s.” The numbers read by the experimenter were written in the verbal format (e.g., “four”). There were twelve randomly presented comparisons (i.e., 1–2, 1–4, 1–6, 1–8, 2–3, 2–9, 3–6, 3–8, 4–9, 6–7, 6–9, 8–9) repeated twice to have the larger number in both locations. For each participant, we calculated the percentage of correct responses as main performance index.

Arabic Digit Comparison

Two digits were horizontally presented, respectively, on the left and right side of the computer screen. The child was asked to indicate the side of the larger digit by pressing the corresponding (left or right) touchpad button. There were 72 randomly presented trials displaying all possible pairs of digits from 1 to 9 twice. The larger number was equally presented in both locations. We calculated the percentage of correct responses as accuracy measure.

Numerosity Estimation

Children verbally estimated the numerosity of a set composed of black squares presented in the center of the screen for 1 s. There were two practice trials (i.e., 2 and 8) and then the numerosities 1, 2, 3, 4, 6, 8, and 10 were randomly displayed four times for a total of 28 trials. For each target numerosity, in half of the sets the item size diminished with increasing numerosity (i.e., equal cumulative surface area) whereas in the other half the item size was constant (i.e., constant item size). We manipulated item size to prevent children from basing their numerical estimates on visual cues instead of focusing on the numerosity of the presented sets. For each participant, we calculated the mean absolute deviation between the estimate and the target number separately for small (≤ 4) and large (> 4) target numerosities. We also computed the individual regression slopes of the estimates as function of target numerosities from 6 to 10 ($M = 0.42$, $SD = 0.49$, range: -0.88 , 1.31). Following Le Corre and Carey (2007)’s classification, children displaying a slope ≥ 0.3 were classified as *verbal mappers* ($n = 30$) whereas other children were classified as *non-mappers* ($n = 17$).

RESULTS

Statistical analyses were conducted using the free software R (R Core Team, 2016) with the following packages: BayesFactor, using default priors (Morey and Rouder, 2015); Hmisc (Harrell et al., 2016); psych (Revelle, 2016); xlsx (Dragulescu, 2014); Rmisc (Hope, 2013); lmSupport (Curtin, 2016); plyr (Wickham, 2011); metafor (Viechtbauer, 2010); car (Fox and Weisberg, 2011); lmtest (Zeileis and Hothorn, 2002); Reshape2 (Wickham, 2007). We report Bayes factors (BF_{10}) expressing the probability of the data given H_1 relative to H_0 (i.e., values larger than 1 are in favor of H_1 , the alternative hypothesis, whereas values smaller than 1 are in favor of H_0 , the null hypothesis). When comparing regression models, we report the Bayes factors (BF) as the ratio of BF_{10} between compared models. If the ratio between BF_{10} of model A and BF_{10} of model B is larger than 1, then there is evidence for model A. Conversely, if the ratio is smaller than one there is evidence for model B. We describe the evidence associated with BFs as “anecdotal” ($1/3 < BF < 3$), “moderate” ($BF < 1/3$ or $BF > 3$), “strong” ($BF < 1/10$ or $BF > 10$), “very strong” ($BF < 1/30$ or $BF > 30$), and “extreme” ($BF < 1/100$ or $BF > 100$) (Jeffreys, 1961; Wagenmakers et al., 2016, 2017). Data and code can be found at https://osf.io/swg8r/?view_only=0fa72144bc1046c99efc0ee258ccf2b9.

We removed those trials with response time below 200 ms (i.e., anticipation) in the computerized tasks: this applied to only one trial in the Arabic digit comparison task. In the numerosity estimation task, we removed absent responses (e.g., “I don’t know”; 3 trials) and trials with estimates above 20 (extreme responses; 23 trials). We ran a Bayesian repeated measures ANOVA on the mean estimate with Target numerosity [1, 2, 3, 4, 6, 8, and 10] and Stimulus set [equal cumulative surface area, constant item size] as within-subjects factors. The model with only Target numerosity yielded the largest evidence compared to the null model ($BF_{10} = 6.39 \times 10^{196}$) and it was superior to

the model also including Stimulus set ($BF_{10} = 8.76 \times 10^{195}$) or the model including the interaction between Target numerosity and Stimulus set ($BF_{10} = 1.31 \times 10^{194}$). This ensured that the estimates did not vary depending on the visual properties of the presented numerical sets (i.e., equal cumulative surface area and constant item size).

The main descriptive statistics of the administered tasks are reported in **Table 1**.

Regression Analyses

We ran two separate regression analyses in order to specifically highlight the contribution of the assessed numerical skills to number words and Arabic digit comparison, respectively. For all the regression models reported in **Tables 2, 3**: residuals were normally distributed (non-significant Shapiro tests, except for Model 1, $p = 0.006$, in **Table 2**; Model 1, $p < 0.001$, and Model 3, $p = 0.017$, in **Table 3**); multicollinearity was absent (i.e., all Variance inflation Factors were lower than 4 for the models with two or more predictors); heteroscedasticity was absent (i.e., non-significant Breusch–Pagan tests, except for Model 1, $p = 0.05$, in **Table 2**); no influential observations were found (i.e., all Cook's distances were below or equal 1).

Number Words Comparison

In the first regression analysis, we used the proportion of correct responses in the number words comparison task (transformed with arcsine square root formula¹; Osborne, 2010) as the outcome variable (**Table 2**). There was extreme evidence for the model including the accuracy in the GaN task (Model 1). Compared to Model 1, there was anecdotal evidence for the models also including PAE in the NL task (Model 2), the absolute deviation for larger numerosities from the numerosity estimation task (Model 3), and all three predictors together (Model 4). We replaced the absolute difference in the numerosity estimation task with a variable coding for the status of verbal mapper (=1) and non-mapper (=0) as proposed by Le Corre and Carey (2007). There was moderate evidence against the model including the status of mapper compared to the model with only GaN performance ($BF = 0.30$), also when

we considered the number word comparison accuracy only for large number words (>4 ; $BF = 0.34$) as in study Le Corre's (2014). Accordingly, verbal mappers and non-mappers displayed a similar accuracy when comparing all number words (Verbal mappers: $M = 87\%$, $SD = 14$; Verbal non-mappers: $M = 85\%$, $SD = 14$; Bayesian t -test: $BF_{10} = 0.32$, moderate evidence) and only large number words (Verbal mappers: $M = 79\%$, $SD = 22$; Verbal non-mappers: $M = 79\%$, $SD = 17$; Bayesian t -test: $BF_{10} = 0.30$, moderate evidence). The same pattern of results emerged when we compared the model with only GaN accuracy with the model also including the linear slope for large numerosities in the numerical estimation task as predictor of all number words comparison ($BF = 0.28$) and large number words comparison ($BF = 0.33$). Finally, there was moderate evidence against the model including age in months and the performance in the GaN compared to the model with only GaN accuracy ($BF = 0.28$), thereby confirming the predominant role of cardinality knowledge. Overall, the results strongly support the relation between cardinality knowledge and number words comparison accuracy, whereas there was no clear evidence for a role of numerosity estimation and spatial mapping abilities.

Arabic Digit Comparison

In the second regression analysis, we used the proportion of correct responses (transformed with arcsine square root formula) in the Arabic digit comparison task as the outcome variable (**Table 3**). There was very strong evidence for the model including GaN accuracy and the PAE in the NL task (Model 2) compared to the model including only the accuracy in the GaN task (Model 1). Conversely, there was anecdotal evidence for the model including the accuracy in the GaN task and numerosity estimation for large numerosities (Model 3) compared to the model with only the accuracy in the GaN task (Model 1). Similarly, there was anecdotal evidence for the model (Model 4) including the absolute deviation for large numerosities in the numerosity estimation task compared to the model including the accuracy in the GaN task and the PAE in the NL task (Model 2).

We also assessed whether performance in the Arabic digit comparison task could be fully accounted for by the ability to compare number words and the accuracy in naming Arabic digits, thereby excluding the influence of spatial mapping. Therefore, in Model 5, we simultaneously included the accuracy in the GaN task, the accuracy in the naming task, and the accuracy in the number words comparison task. There was extreme evidence for the inclusion of the accuracy in naming and in the comparison of number words (Model 5) compared to the model with only the accuracy in the GaN task (Model 1). In Model 6, we also entered the PAE in the NL task. We found moderate evidence for the model also including the PAE in the NL task, thereby suggesting a specific contribution of spatial mapping to the understanding of magnitude relation between Arabic digits (Sella et al., 2017). Accordingly, spatial mappers were more accurate in comparing Arabic digits compared to non-mappers (Spatial mappers: $M = 90\%$, $SD = 14$; Spatial non-mappers: $M = 60\%$, $SD = 13$; Bayesian t -test: $BF_{10} = 403543$).

¹The same pattern of results emerged when regression models were run on the proportion of correct responses in the number words comparison task and in the Arabic digit comparison task.

TABLE 1 | Descriptive statistics for the administered numerical tasks.

Task	<i>M</i>	<i>SD</i>	95% CI
Naming (1–9) (% of correct responses)	79	30	[70–88]
Give-a-Number (% of correct responses)	88	21	[81–94]
NL task (PAE)	20	12	[16–23]
Numerosities estimation (absolute difference)			
- Small numerosities (≤ 4)	0.22	0.3	[0.13–0.3]
- Large numerosities (> 4)	2.09	0.89	[1.82–2.35]
Arabic digit comparison (% of correct responses)	82	19	[77–88]
Number words comparison (% of correct responses)	86	14	[82–91]

TABLE 2 | Summary of the regression models with proportion of correct responses (arcsine transformed) in the number word comparison task as outcome variable.

Model	Measures	<i>B</i>	95% CI	Bayes factor (BF ₁₀) for the comparison with the null model	<i>R</i> ²	Bayes factor for model comparison
1	GaN (% correct)	0.006	[0.003 0.009]	144	0.27	
2	GaN (% correct)	0.004	[0.001 0.008]	107	0.31	BF ₁₀ model 2/BF ₁₀ model 1 = 0.74
3	NL (PAE)	−0.005	[−0.011 0.002]	160	0.32	BF ₁₀ model 3/BF ₁₀ model 1 = 1.11
	GaN (% correct)	0.005	[0.002 0.008]			
4	Numerosity estimation large numerosities (absolute difference)	−0.066	[−0.139 0.007]	91	0.35	BF ₁₀ model 4/BF ₁₀ model 1 = 0.63
	GaN (% correct)	0.004	[0.0003 0.0075]			
	NL (PAE)	−0.004	[−0.010 0.003]			
	Numerosity estimation large numerosities (absolute difference)	−0.057	[−0.131 0.018]			

TABLE 3 | Summary of the regression models with proportion of correct responses (arcsine transformed) in the Arabic digit comparison task as outcome variable.

Model	Measures	<i>B</i>	95% CI	Bayes factor (BF ₁₀) for the comparison with the null model	<i>R</i> ²	Bayes factor for model comparison
1	GaN (% correct)	0.007	[0.004 0.011]	395	0.31	
2	GaN (% correct)	0.004	[−0.0001 0.007]	14338	0.46	BF ₁₀ model 2/BF ₁₀ model 1 = 36
3	NL (PAE)	−0.011	[−0.018 −0.005]	674	0.37	BF ₁₀ model 3/BF ₁₀ model 1 = 1.7
	GaN (% correct)	0.006	[0.003 0.009]			
4	Numerosity estimation large numerosities (absolute difference)	−0.085	[−0.167 −0.004]	10731	0.49	BF ₁₀ model 4/BF ₁₀ model 2 = 0.75
	GaN (% correct)	0.003	[−0.001 0.007]			
	NL (PAE)	−0.010	[−0.017 −0.004]			
5	Numerosity estimation large numerosities (absolute difference)	−0.059	[−0.135 0.016]	4.11 × 10 ⁷	0.66	BF ₁₀ model 5/BF ₁₀ model 1 = 104155
	GaN (% correct)	0.001	[−0.002 0.004]			
	Naming (% correct)	0.002	[−0.0004 0.004]			
6	Number words comparison (% correct)	0.013	[0.008 0.017]	2.3 × 10 ⁸	0.72	BF ₁₀ model 6/BF ₁₀ model 5 = 5.61
	GaN (% correct)	0.001	[−0.003 0.004]			
	Naming (% correct)	−0.001	[−0.003 0.002]			
	Number words comparison (% correct)	0.012	[0.008 0.016]			
	NL (PAE)	−0.009	[−0.015 −0.002]			

The absolute difference for small numerosities in the numerosity estimation task was never a relevant predictor when entered in the previous regression models (all BF_s < 1).

DISCUSSION

In the present study, we investigated the specific role of numerosity estimation and spatial mapping of numbers in the ability to compare auditorily presented number words and visually presented Arabic digits. Previous studies suggested that the ability to compare number words might be associated

with numerosity estimation after controlling for cardinality knowledge (Le Corre, 2014). Similarly, the comparison of Arabic digits has been related to the ability to spatially map numbers on the visual line (Sella et al., 2017). Here, the comparison of number words related to cardinality knowledge but not to numerical estimation or spatial mapping accuracy. Children who knew that later number words in the counting list are associated with larger numerical quantities (i.e., verbal mappers) or were more accurate in mapping numbers on the visual line did not show a better performance in choosing the larger between two number words. Conversely, the ability to spatially map numbers strongly related to the comparison of visually presented Arabic digits. Crucially,

we found moderate evidence for the relation between spatial mapping and Arabic number comparison even after controlling for the accuracy in reading Arabic digits and comparing number words, thereby addressing the potential caveat that the task might be transformed into verbal comparison after transcoding the digits into number words. These results suggest that the comparison of Arabic digits entails a specific spatial component that is captured by the accuracy of spatial mapping. In this regard, we have previously suggested that the spatial arrangement of digits along the line may scaffold the representation of exact numerical magnitude (Sella et al., 2017).

Previous studies have shown that the acquisition of the cardinality principle does not imply a mapping between exact magnitude and number words but rather entails the understanding that the last recited number word denotes the cardinality of the counted set (Davidson et al., 2012; Le Corre, 2014). This view is supported by the finding that some children who have acquired the cardinality principle, as measured by the GaN task, fail in choosing the larger between two number words within their counting range (Le Corre, 2014). It has been proposed that the understanding that later number words in the counting list are associated with large numerical quantities (i.e., the later-greater principle), or, more broadly, the precision of numerosity estimation (i.e., ANS-to-word mapping) might lead children to infer the numerical magnitude of number words. Nevertheless, this view was neither supported nor discarded by the results of the present study. A replication with a larger sample size would disentangle whether the estimation of large numerical quantities is actually related to the ability to compare number words. Conversely, spatial mapping ability was clearly related to understanding the exact magnitude of numerals, as measured by the Arabic digit comparison task. In this vein, the numerical magnitude of a digit can be conveyed by its spatial position on the line and with respect to the positions of other digits, conceivably through a symbol-to-symbol relation (Nieder, 2005; Vogel et al., 2014; Reynvoet and Sasanguie, 2016). The correlational nature of our study prevents us from inferring any casual direction between spatial mapping of numbers and magnitude understanding. Nevertheless, there is evidence that training spatial mapping of numbers leads to better performance in comparing Arabic digits (Siegler and Ramani, 2009; Ramani et al., 2012), thereby supporting the role of the “spatial mapping principle” (Sella et al., 2017) in the acquisition of exact numerical meaning of symbolic numbers.

The results of the present study suggest that children rely on multiple numerical skills and representations, which are differently weighted depending on numerical format. The presentation format plausibly leads children to rely on distinctive representations and strategies when comparing symbolic numbers. In the case of number words, the verbal format might lead children to rely on a verbal mechanism, such as counting. In this regard, it is worth noting that the comparison of each pair of number words was followed by the presentation of the corresponding numerosities in the current experimental paradigm (following Le Corre, 2014). In addition to providing visual feedback on the choice, this is likely to have triggered an enumeration strategy, even though children were not allowed to

count the elements in the set but were moved immediately to the next trial. Conversely, the visual presentation of Arabic digits may have triggered a “number line” representation to choose the larger digit based on its spatial position.

More broadly, young children progressively integrate multiple representations of numerical information (verbal, visual, and analogical) and learn to switch from one representation to another (Dehaene and Cohen, 1995; Kucian and Kaufmann, 2009). It has been suggested that children first map number words to numerical sets, then map Arabic digits to numerical sets, and finally associate number words to Arabic digits (Benoit et al., 2013). Conversely, others have found that children first create an association between number words and the corresponding numerical quantities, then associate number words to Arabic numerals (Hurst et al., 2016). Similarly, preschool children first learn the cardinal meaning of number words, then to read Arabic digits, and finally learn the cardinal meaning of numerals and how to order them (Knudsen et al., 2015). Overall, this reveals a complex scenario in which children build connections between different representations of numbers in a relatively short time window. The integration of different representations of numbers is likely to be heavily influenced by individual experience that children have with numbers. For example, Arabic numerals might be introduced at different times across a sample of preschool children, which would clearly affect their understanding of their numerical meaning. Therefore, it would not be surprising to observe variability and divergent developmental patterns across different studies. Future research may describe the different developmental patterns associated with the integration of multiple representations of numbers and highlight the more efficient ways to achieve full understanding of the magnitudes associated with symbolic numbers. This kind of evidence would be extremely valuable for cognitive scientists and for educators interested in improving children’s early numerical skills.

CONCLUSION

Preschool children use different numerical skills and representations depending on the presentation format to compare the numerical magnitude of symbolic numbers. The results of the present study suggest that the comparison of number words relates to cardinality knowledge whereas the comparison of Arabic numerals specifically relates to the spatial mapping of numbers. This finding supports the hypothesis that a spatial mapping principle scaffolds the acquisition of symbolic number knowledge.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the national ethics guidelines for psychologists [Codice Deontologico degli Psicologi Italiani] with written informed consent obtained from parents or legal guardians. All parents or legal guardians gave written informed

consent in accordance with the Declaration of Helsinki. The protocol was approved by the Ethics Committee for Psychology Research of the University of Padova.

AUTHOR CONTRIBUTIONS

FS, DL and MZ designed the study. FS supervised data collection, performed statistical analyses, and drafted the manuscript. DL and MZ provided critical revisions of the article. All authors approved the final version of the manuscript.

REFERENCES

- Benoit, L., Lehalle, H., Molina, M., Tijus, C., and Jouen, F. (2013). Young children's mapping between arrays, number words, and digits. *Cognition* 129, 95–101. doi: 10.1016/j.cognition.2013.06.005
- Berteletti, I., Lucangeli, D., Piazza, M., Dehaene, S., and Zorzi, M. (2010). Numerical estimation in preschoolers. *Dev. Psychol.* 46, 545–551. doi: 10.1037/a0017887
- Booth, J. L., and Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Dev. Psychol.* 42, 189–201. doi: 10.1037/0012-1649.41.6.189
- Carey, S. (2001). Cognitive foundations of arithmetic: evolution and ontogenesis. *Mind Lang.* 16, 37–55. doi: 10.1111/1468-0017.00155
- Curtin, J. (2016). *lmSupport: Support for Linear Models*. Available at: <https://cran.r-project.org/package=lmSupport>
- Davidson, K., Eng, K., and Barner, D. (2012). Does learning to count involve a semantic induction? *Cognition* 123, 162–173. doi: 10.1016/j.cognition.2011.12.013
- De Hevia, M. D., Girelli, L., Addabbo, M., and Macchi Cassia, V. (2014). Human infants' preference for left-to-right oriented increasing numerical sequences. *PLoS One* 9:e96412. doi: 10.1371/journal.pone.0096412
- de Hevia, M. D., Vallar, G., and Girelli, L. (2008). Visualizing numbers in the mind's eye: the role of visuo-spatial processes in numerical abilities. *Neurosci. Biobehav. Rev.* 32, 1361–1372. doi: 10.1016/j.neubiorev.2008.05.015
- Dehaene, S. (1992). Varieties of numerical abilities. *Cognition* 44, 1–42. doi: 10.1016/0010-0277(92)90049-N
- Dehaene, S., Bossini, S., and Giraux, P. (1993). The mental representation of parity and number magnitude. *J. Exp. Psychol. Gen.* 122, 371–396. doi: 10.1037/0096-3445.122.3.371
- Dehaene, S., and Cohen, L. (1995). Towards an anatomical and functional model of number processing. *Math. Cogn.* 1, 83–120.
- Dragulescu, A. A. (2014). *xlsx: Read, Write, Format Excel 2007 and Excel 97/2000/XP/2003 Files*. Available at: <https://cran.r-project.org/package=xlsx>
- Fischer, M. H., Castel, A. D., Dodd, M. D., and Pratt, J. (2003). Perceiving numbers causes spatial shifts of attention. *Nat. Neurosci.* 6, 555–556. doi: 10.1038/nn1066
- Fox, J., and Weisberg, S. (2011). *An Companion to Applied Regression*, 2nd Edn. Thousand Oaks, CA: Sage.
- Gebuis, T., and Reynvoet, B. (2011). Generating nonsymbolic number stimuli. *Behav. Res. Methods* 43, 981–986. doi: 10.3758/s13428-011-0097-5
- Gelman, R., and Gallistel, C. R. (1978). *The Child's Understanding of Number. Educational Researcher*, Vol. 8. Cambridge, MA: Harvard University Press.
- Harrell, F. E. Jr., Dupont, C., and Al, E. (2016). *Hmisc: Harrell Miscellaneous*. Available at: <https://cran.r-project.org/package=Hmisc>
- Hope, R. M. (2013). *Rmisc: Ryan Miscellaneous*. Available at: <https://cran.r-project.org/package=Rmisc>
- Hubbard, E. M., Piazza, M., Pinel, P., and Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nat. Rev. Neurosci.* 6, 435–448. doi: 10.1038/nrn1684
- Hurst, M., Anderson, U., and Cordes, S. (2016). Mapping among number words, numerals, and non-symbolic quantities in preschoolers. *J. Cogn. Dev.* 18, 41–62. doi: 10.1080/15248372.2016.1228653
- Jeffreys, H. (1961). *Theory of Probability*, 3rd Edn. Oxford: Oxford University Press.

FUNDING

This study was supported by the University of Padova (Strategic Grant “NEURAT” to MZ).

ACKNOWLEDGMENTS

The authors wish to thank the children and their parents for participating in the present study, as well as Sara Borsatto for her help in collecting data.

- Jiménez Lira, C., Carver, M., Douglas, H., and LeFevre, J. A. (2017). The integration of symbolic and non-symbolic representations of exact quantity in preschool children. *Cognition* 166, 382–397. doi: 10.1016/j.cognition.2017.05.033
- Knudsen, B., Fischer, M. H., Henning, A., and Aschersleben, G. (2015). The development of Arabic digit knowledge in 4- to 7-year-old children. *J. Numer. Cogn.* 1, 21–37. doi: 10.5964/jnc.v1i1.4
- Kucian, K., and Kaufmann, L. (2009). A developmental model of number representation. *Behav. Brain Sci.* 32, 340–341. doi: 10.1017/S0140525X09990069
- Le Corre, M. (2014). Children acquire the later-greater principle after the cardinal principle. *Br. J. Dev. Psychol.* 32, 163–177. doi: 10.1111/bjdp.12029
- Le Corre, M., and Carey, S. (2007). One, two, three, four, nothing more: an investigation of the conceptual sources of the verbal counting principles. *Cognition* 105, 395–438. doi: 10.1016/j.cognition.2006.10.005
- Morey, R. D., and Rouder, J. N. (2015). *BayesFactor: Computation of Bayes Factors for Common Designs. R Package Version 0.9.12-12*.
- Nieder, A. (2005). Counting on neurons: the neurobiology of numerical competence. *Nat. Rev. Neurosci.* 6, 177–190. doi: 10.1038/nrn1626
- Nuerk, H. C., Patro, K., Cress, U., Schild, U., Friedrich, C. K., and Gobel, S. M. (2015). How space-number associations may be created in preliterate children: six distinct mechanisms. *Front. Psychol.* 6:215. doi: 10.3389/fpsyg.2015.00215
- Osborne, J. W. (2010). Improving your data transformations?: applying the Box-Cox transformation. *Pract. Assess. Res. Eval.* 15, 1–9.
- Patro, K., Fischer, U., Nuerk, H. C., and Cress, U. (2016). How to rapidly construct a spatial-numerical representation in preliterate children (at least temporarily). *Dev. Sci.* 19, 126–144. doi: 10.1111/desc.12296
- Patro, K., and Haman, M. (2012). The spatial-numerical congruity effect in preschoolers. *J. Exp. Child Psychol.* 111, 534–542. doi: 10.1016/j.jecp.2011.09.006
- R Core Team (2016). *R Development Core Team. R: A Language and Environment for Statistical Computing*. Available at: <https://www.r-project.org/>
- Ramani, G. B., Siegler, R. S., and Hitti, A. (2012). Taking it to the classroom: number board games as a small group learning activity. *J. Educ. Psychol.* 104, 661–672. doi: 10.1037/a0028995
- Restle, F. (1970). Speed of adding and comparing numbers. *J. Exp. Psychol.* 83(2 Pt 1), 274–278. doi: 10.1037/h0028573
- Revelle, W. (2016). *psych: Procedures for Psychological, Psychometric, and Personality Research. Evanston, Illinois*. Available at: <http://cran.r-project.org/package=psych>
- Reynvoet, B., and Sasanguie, D. (2016). The symbol grounding problem revisited: a thorough evaluation of the ANS mapping account and the proposal of an alternative account based on symbol-symbol associations. *Front. Psychol.* 7:1581. doi: 10.3389/fpsyg.2016.01581
- Rugani, R., Vallortigara, G., Priftis, K., and Regolin, L. (2015). Number-space mapping in the newborn chick resembles humans mental number line. *Science* 347, 534–536. doi: 10.1080/135765004420
- Sarnecka, B. W. (2015). Learning to represent exact numbers. *Synthese* 1–18. doi: 10.1007/s11229-015-0854-6
- Sarnecka, B. W., and Carey, S. (2008). How counting represents number: what children must learn and when they learn it. *Cognition* 108, 662–674. doi: 10.1016/j.cognition.2008.05.007
- Sella, F., Berteletti, I., Lucangeli, D., and Zorzi, M. (2017). Preschool children use space, rather than counting, to infer the numerical magnitude of digits: evidence

- for a spatial mapping principle. *Cognition* 158, 56–67. doi: 10.1016/j.cognition.2016.10.010
- Siegler, R. S., and Booth, J. L. (2004). Development of numerical estimation in young children. *Child Dev.* 75, 428–444. doi: 10.1111/j.1467-8624.2004.00684.x
- Siegler, R. S., and Opfer, J. E. (2003). The development of numerical estimation: evidence for multiple representations of numerical quantity. *Psychol. Sci.* 14, 237–243. doi: 10.1111/1467-9280.02438
- Siegler, R. S., and Ramani, G. B. (2009). Playing linear number board games—but not circular ones—improves low-income preschoolers' numerical understanding. *J. Educ. Psychol.* 101, 545–560. doi: 10.1037/a0014239
- Siegler, R. S., Thompson, C. A., and Opfer, J. E. (2009). The logarithmic-to-linear shift: one learning sequence, many tasks, many time scales. *Mind Brain Educ.* 3, 143–150. doi: 10.1111/j.1751-228X.2009.01064.x
- Umiltà, C., Priftis, K., and Zorzi, M. (2009). The spatial representation of numbers: evidence from neglect and pseudoneglect. *Exp. Brain Res.* 192, 561–569. doi: 10.1007/s00221-008-1623-2
- Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* 36, 1–48. doi: 10.18637/jss.v036.i03
- Vogel, S. E., Remark, A., and Ansari, D. (2014). Differential processing of symbolic numerical magnitude and order in first-grade children. *J. Exp. Child Psychol.* 129, 26–39. doi: 10.1016/j.jecp.2014.07.010
- Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., et al. (2017). Bayesian inference for psychology. Part II: example applications with JASP. *Psychon. Bull. Rev.* 25, 58–76. doi: 10.3758/s13423-017-1323-7
- Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., et al. (2016). Bayesian inference for psychology. Part I: theoretical advantages and practical ramifications. *Psychon. Bull. Rev.* 25, 169–176. doi: 10.3758/s13423-017-1343-3
- Wickham, H. (2007). Reshaping data with the reshape package. *J. Stat. Softw.* 12, 1–20. doi: 10.3978/j.issn.2305-5839.2016.01.33
- Wickham, H. (2011). The split-apply-combine strategy for data analysis. *J. Stat. Softw.* 40, 1–29. doi: 10.18637/jss.v040.i01
- Wynn, K. (1990). Children's understanding of counting. *Cognition* 36, 155–193. doi: 10.1016/0010-0277(90)90003-3
- Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cogn. Psychol.* 24, 220–251. doi: 10.1016/0010-0285(92)90008-P
- Zeileis, A., and Hothorn, T. (2002). Diagnostic checking in regression relationships. *R News* 2, 7–10.
- Zorzi, M., Bonato, M., Treccani, B., Scalabrini, G., Marensi, R., and Priftis, K. (2012). Neglect impairs explicit processing of the mental number line. *Front. Hum. Neurosci.* 6:125. doi: 10.3389/fnhum.2012.00125
- Zorzi, M., Priftis, K., and Umiltà, C. (2002). Neglect disrupts the mental number line. *Nature* 417, 138–140. doi: 10.1038/417138a

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Sella, Lucangeli and Zorzi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Implicit and Explicit Number-Space Associations Differentially Relate to Interference Control in Young Adults With ADHD

Carrie Georges^{1*}, Danielle Hoffmann² and Christine Schiltz¹

¹ Institute of Cognitive Science and Assessment, Research Unit Education, Culture, Cognition and Society, Faculty of Language and Literature, Humanities, Arts and Education, University of Luxembourg, Luxembourg, Luxembourg,

² Luxembourg Centre for Educational Testing, Faculty of Language and Literature, Humanities, Arts and Education, University of Luxembourg, Luxembourg, Luxembourg

OPEN ACCESS

Edited by:

Maciej Haman,
University of Warsaw, Poland

Reviewed by:

Thomas Dresler,
Universität Tübingen, Germany
Malgorzata Gut,
Nicolaus Copernicus University
in Toruń, Poland

*Correspondence:

Carrie Georges
carrie.georges@uni.lu

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 13 October 2017

Accepted: 01 May 2018

Published: 24 May 2018

Citation:

Georges C, Hoffmann D and
Schiltz C (2018) Implicit and Explicit
Number-Space Associations
Differentially Relate to Interference
Control in Young Adults With ADHD.
Front. Psychol. 9:775.
doi: 10.3389/fpsyg.2018.00775

Behavioral evidence for the link between numerical and spatial representations comes from the spatial-numerical association of response codes (SNARC) effect, consisting in faster reaction times to small/large numbers with the left/right hand respectively. The SNARC effect is, however, characterized by considerable intra- and inter-individual variability. It depends not only on the explicit or implicit nature of the numerical task, but also relates to interference control. To determine whether the prevalence of the latter relation in the elderly could be ascribed to younger individuals' ceiling performances on executive control tasks, we determined whether the SNARC effect related to Stroop and/or Flanker effects in 26 young adults with ADHD. We observed a divergent pattern of correlation depending on the type of numerical task used to assess the SNARC effect and the type of interference control measure involved in number-space associations. Namely, stronger number-space associations during parity judgments involving implicit magnitude processing related to weaker interference control in the Stroop but not Flanker task. Conversely, stronger number-space associations during explicit magnitude classifications tended to be associated with better interference control in the Flanker but not Stroop paradigm. The association of stronger parity and magnitude SNARC effects with weaker and better interference control respectively indicates that different mechanisms underlie these relations. Activation of the magnitude-associated spatial code is irrelevant and potentially interferes with parity judgments, but in contrast assists explicit magnitude classifications. Altogether, the present study confirms the contribution of interference control to number-space associations also in young adults. It suggests that magnitude-associated spatial codes in implicit and explicit tasks are monitored by different interference control mechanisms, thereby explaining task-related intra-individual differences in number-space associations.

Keywords: SNARC effect, magnitude processing, interference control, Stroop effect, Flanker effect, individual differences

INTRODUCTION

Numbers and space are closely associated in the human mind (e.g., Dehaene and Brannon, 2011). The most extensively studied and replicated behavioral evidence for this association is without a doubt the spatial-numerical association of response codes (SNARC) effect (Dehaene et al., 1993). It describes the observation that individuals from Western societies are typically faster on their left/right hand-side for relatively small/large numbers respectively, when doing binary classifications on numbers. The SNARC effect was first documented in an experiment where numerical magnitude information was task-relevant (termed “the magnitude SNARC effect”) in that individuals judged whether a centrally displayed number was smaller or larger than a given standard (Dehaene et al., 1990). Subsequent experiments, however, demonstrated that numerical magnitude does not need to be task-relevant to observe the SNARC effect, since it was also evidenced during parity judgments (termed “the parity SNARC effect”; e.g., Dehaene et al., 1993).

Three spatial coding mechanisms were proposed to account for spatial-numerical interactions, including a visuospatial, verbal-spatial, and working memory (WM) account (for a review, see e.g., Fischer and Shaki, 2014). According to the dominant and most traditional *visuospatial* account, numbers are mentally represented along a continuous left-to-right-oriented spatial representational medium, also known as the mental number line (MNL), with small/large numbers located on its left/right respectively, at least in Western societies (Moyer and Landauer, 1967; Restle, 1970; Dehaene et al., 1993). An alternative view suggests that number-space associations arise from categorical *verbal-spatial* coding. The latter account is based on the polarity correspondence principle by Proctor and Cho (2006) and assumes that the SNARC effect results from the polar correspondence between the verbal categorical concepts “small” and “left” (both assigned to the same polarity) as well as “large” and “right” (both assigned to the opposing polarity). A final explanation for the link between numbers and space was provided by Fias et al. (2011), who argued that spatial-numerical interactions are task-specific associations established within WM (see also van Dijck and Fias, 2011; Abrahamse et al., 2016; Fias and van Dijck, 2016). More concretely, task-relevant numerical magnitudes are temporarily activated in their canonical order within a horizontal left-to-right oriented spatial sequence in WM. Spatial-numerical interactions then result from internal shifts of spatial attention within this encoded numerical sequence, with positions from the beginning/end of the sequence eliciting faster left-/right-sided responses respectively.

Inter-Individual Differences in Number-Space Associations

The strength of number-space associations considerably varies between individuals. For instance, variability is explained by inter-individual differences in mathematical skills. Participants with lower arithmetic performances featured stronger number-space associations in the parity judgment task (e.g., Georges

et al., 2017; but see Cipora and Nuerk, 2013). Similarly, more pronounced parity SNARC effects were observed in humanities students with than without math difficulties (Hoffmann et al., 2014a), while the weakest number-space associations were evidenced in math professionals (Cipora et al., 2016). The parity SNARC effect also depends on math anxiety, with more anxious individuals displaying stronger number-space associations (Georges et al., 2016). Furthermore, it was shown to increase with age (Hoffmann et al., 2014b; Ninaus et al., 2017).

In addition to this, inter-individual variability in the parity SNARC effect has recently been shown to relate to differences in inhibitory control as indexed by the Stroop effect (Hoffmann et al., 2014b; for a review on the Stroop effect, see MacLeod, 1991; see also Stroop, 1935). Participants with weaker interference control in the Stroop paradigm featured stronger number-space associations in the parity judgment task. The relation between number-space associations during parity judgments and inhibitory control might be explained by the need to inhibit numerical magnitude and its associated spatial code to accurately respond based on the number's parity status. It should, however, be noted that the relation between weaker interference control in the Stroop task and stronger parity SNARC effects was most pronounced in the elderly. It did not reach significance in young healthy individuals, which the authors ascribed to their near ceiling performances on the Stroop task.

Intra-Individual Differences in Number-Space Associations

Apart from inter-individual differences in the SNARC effect, number-space associations also vary intra-individually depending on the number processing task. For instance, Georges et al. (2017) observed no significant relation between the SNARC effects in a parity judgment and magnitude classification task (at least at the sample level – positive and negative correlations were evidenced in individuals with object and spatial visualization styles respectively). Moreover, verbal and visuospatial WM load selectively abolished the parity and magnitude SNARC effects respectively (Herrera et al., 2008; van Dijck et al., 2009). In addition, hemi-neglect patients were shown to display regular number-space associations in the parity judgment task, where access to numerical magnitude is implicit, but featured an atypical SNARC effect in the explicit magnitude classification task (Priftis et al., 2006; Zorzi et al., 2012). The SNARC effects in implicit and explicit tasks were also shown to associate with different cognitive factors. Namely, only the magnitude SNARC effect related to inter-individual differences in visualization cognitive styles (Georges et al., 2017; see also Kozhevnikov et al., 2005; Chabris et al., 2006). Furthermore, the relation between weaker arithmetic performances and stronger SNARC effects during parity judgments (e.g., Hoffmann et al., 2014a; Georges et al., 2017; but see Cipora and Nuerk, 2013) was not observed for number-space associations in the explicit magnitude classification task. Altogether, these findings suggest that numbers might be associated with qualitatively different spatial codes depending on the implicit or explicit nature of the numerical processing task.

Interference Control and Inter-/Intra-Individual Differences in Number-Space Associations

The present study aimed to (a) replicate the previously reported relationship between implicit number-space associations and inhibitory control (Hoffmann et al., 2014b) and (b) investigate whether this relationship extends to explicit magnitude processing.

While the “parity SNARC-Stroop” relation was significant in a group composed of young and elderly healthy participants, it was mainly driven by the elderly and did not reach significance in the young subgroup (Hoffmann et al., 2014b). We reasoned that this result pattern might be caused by the fact that young healthy adults achieved near ceiling performances on the Stroop task. In the current study, we therefore focussed on young individuals featuring atypical inhibitory control and included only participants formally diagnosed with ADHD and/or displaying symptoms consistent with ADHD according to the Adult ADHD Self-Report Scale (Kessler et al., 2005). These people not only feature weaker interference control (Walker et al., 2000; Rapport et al., 2001; Lansbergen et al., 2007), but their deficits are also highly variable (Lovejoy et al., 1999; Sergeant et al., 2002; Seidman, 2006). Such inter-individual variability in inhibitory control deficits should increase the statistical power of detecting significant relations with other continuous variables (e.g., Goodwin and Leech, 2006). This enables us to verify whether the previously reported null relation between the parity SNARC and Stroop effects in the younger healthy individuals (Hoffmann et al., 2014b) can indeed be ascribed to their near ceiling performances on the Stroop task. Finding evidence for a significant association between number-space associations in the parity judgment task and interference control in the Stroop paradigm in a relatively younger population would considerably strengthen the critical involvement of inhibitory control mechanisms in the spatial coding processes underlying the parity SNARC effect.

In addition to interference control in the Stroop task, the present study also determined whether executive control processes in the arrowhead version of the Flanker task (e.g., Stins et al., 2004; Davelaar and Stevens, 2009; for the original version, see Eriksen and Eriksen, 1974) might relate to inter-individual differences in number-space associations during parity judgments. Even though conflict occurs in both the Stroop and Flanker paradigms, its nature and processing likely differ depending on the executive control task. For instance, while elderly people were shown to display weaker interference control in the Stroop task than young adults (West and Alain, 2000; Van der Elst et al., 2006), inhibitory control in the Flanker task did not differ between younger and older participants (Falkenstein et al., 2001; Nieuwenhuis et al., 2002). Moreover, heritability of interference control was evidenced in the Stroop but not the Flanker task (Stins et al., 2004). In addition, interference control in the Stroop but not the Flanker task was related to WM capacity (Stins et al., 2005). Furthermore, relations could be evidenced neither between the time needed for conflict resolution nor between the interference scores in the Stroop

and Flanker tasks (Stins et al., 2005). Conflict processing in the Flanker task was shown to relate to the activation of the right dorsolateral prefrontal cortex and the insula (Zhu et al., 2010; Zmigrod et al., 2016). Conversely, neural responses reflecting the Stroop effect were measured in a broader network including not only the right dorsolateral prefrontal cortex, but also the posterior parietal, anterior cingulate and left premotor cortices (van Veen and Carter, 2005; Melcher and Gruber, 2009; Kim et al., 2011; for a meta-analysis, see Nee et al., 2007). These findings thus suggest that Stroop and Flanker effects likely reflect qualitatively different executive control processes. Consequently, contrasting their relations with number-space associations will allow for a better understanding of the specific inhibitory control mechanisms contributing to spatial-numerical interactions.

In a second step, we aimed to assess the relations between the SNARC effect during explicit magnitude classifications and inhibitory control indexed by the Stroop and Flanker effects, since number-space associations were previously shown to vary intra-individually depending on the implicit or explicit nature of the number processing task (van Dijck et al., 2009; Georges et al., 2017). This will inform us about the involvement of inhibitory control processes in the spatial coding processes underlying the magnitude SNARC effect and as such their role in intra-individual differences in number-space associations.

MATERIALS AND METHODS

This study was reviewed and approved by the Ethics Review Panel (ERP) of the University of Luxembourg. All participants gave written informed consent and received a small monetary reward for their participation.

Participants

The study was advertised to the university students via their email addresses. Students could take part in the study if they were formally diagnosed with ADHD (Attention-Deficit/Hyperactivity Disorder) and/or if they considered themselves as being easily distracted and unable to concentrate. A total of 42 students signed up for the study, of which 5 had a formal diagnosis of ADHD. Participants had various backgrounds with different mother tongues (e.g., English, Finnish, French, German, Greek, Russian, Spanish, etc.) and their study fields ranged from mathematics and physics over law to humanities. None of the participants suffered from any comorbid learning disabilities such as dyslexia or dyscalculia.

Procedure and Tasks

Before the start of the experiment, the 42 students that had signed up for the study completed the 6-item version of the World Health Organization Adult ADHD Self-Report Scale V 1.1 (ASRS) symptom checklist (Kessler et al., 2005; for psychometric properties, see Adler et al., 2006; Matza et al., 2011). This was to ensure that individuals not formally diagnosed with ADHD displayed symptoms consistent with this disorder. Participants without a formal diagnosis of ADHD that did not feature ADHD traits according to this self-report scale were excluded prior to the

start of the study. This reduced the study sample to a total of 35 participants.

These participants completed the experimental tasks during two testing sessions that were run on separate days with an upper limit of 1 week apart. Following standard practice in individual differences research (e.g., Carlson and Moses, 2001), all participants performed the tests in the same order and trial sequences were identical for all participants in every task. On the first testing day, participants completed the speeded matching-to-sample task, the parity judgment task, the magnitude classification task and the Flanker task. These computerized tasks were programmed in E-prime (Version 1.2 or 2.0.8.79) and administered on a Windows computer. The classical verbal paper-and-pencil version of the Stroop task was implemented on the second testing day.

Prior to data analysis, 4 students were excluded from the sample since they did not complete all the tests. After removal of these participants, outliers were identified for each of the measures described below. A total of 5 participants had to be removed, since their performances fell 2.5 standard deviations (*SD*) below or above the mean group performances on at least one of the measures. All statistical analyses were thus conducted on data obtained from 26 individuals.

Parity Judgment and Magnitude Classification Tasks

The **parity judgment task** (adapted from Dehaene et al., 1993; see also Georges, 2017; see **Figure 1A**) was administered to determine number-space associations in a task with **implicit numerical magnitude processing**. The experiment consisted of 288 experimental trials divided equally across two blocks. Each experimental trial started with an empty black-bordered square ($6.87^\circ \times 6.87^\circ$) on a white background. After 300 ms, one of eight possible stimuli (Arabic digits: 1, 2, 3, 4, 6, 7, 8, or 9; color: black; font: Arial; point size: 64) appeared in the center of the black-bordered square and remained until response. The inter-trial interval consisted of a blank screen of 1300 ms. In the first block, participants judged as quickly as possible whether the presented number was odd/even by pressing the “A”/“L” key on a QWERTZ keyboard respectively. This stimulus-response mapping was reversed for all participants in the second block. Each target number was displayed 18 times per block. The sequence in which the target stimuli appeared was pseudo-randomized in a way that no target number could appear twice in a row, and the correct response could not be on the same side more than three times consecutively. Each block started with 12–20 training trials, depending on response accuracy. Participants were given a small break half-way through each block.

The **magnitude classification task** (adapted from Bull et al., 2005; van Galen and Reitsma, 2008; see also Georges, 2017; see **Figure 1A**) was administered to determine number-space associations in a task with **explicit numerical magnitude processing**. The experiment was identical to the parity judgment task with the exception that it only consisted of 144¹ trials and

that participants had to judge whether the centrally presented single Arabic number was smaller/larger than five by pressing the “A”/“L” key respectively in the first block. This stimulus-response mapping was again reversed for all participants in the second block.

Data from the training sessions was not analyzed (for comparable data analysis, see Georges et al., 2017). The mean error rate on experimental trials was 2.52 and 2.56% in the parity judgment and magnitude classification task respectively [$F(1,25) = 0.006$; $p = 0.94$; $\eta_p^2 = 0.00$]. Errors were not further analyzed. Reaction times (RTs) shorter or longer than 2.5 *SD* from the individual mean were considered as outliers and discarded prior to data analysis (2.86 and 3.19% of all correct trials in the parity judgment and magnitude classification task respectively, $F(1,25) = 1.55$; $p = 0.23$; $\eta_p^2 = 0.06$).

SNARC effect regression slopes were computed using the individual regression equations method suggested by Fias et al. (1996). First, RTs were averaged separately for each number and each response side for every participant. Individual RT differences (dRTs) were then calculated by subtracting for each number the mean left-sided RT from the mean right-sided RT. The resulting dRTs were subsequently submitted to a regression analysis, using number magnitude as predictor variable. Unstandardized SNARC regression slopes were taken as a measure of the strength of the SNARC effect in terms of the inclination of the regression lines. Negative regression weights reflected SNARC effects in the expected direction (faster left-/right-sided RTs for small/large numbers respectively) with more negative regression slopes corresponding to stronger number-space associations.

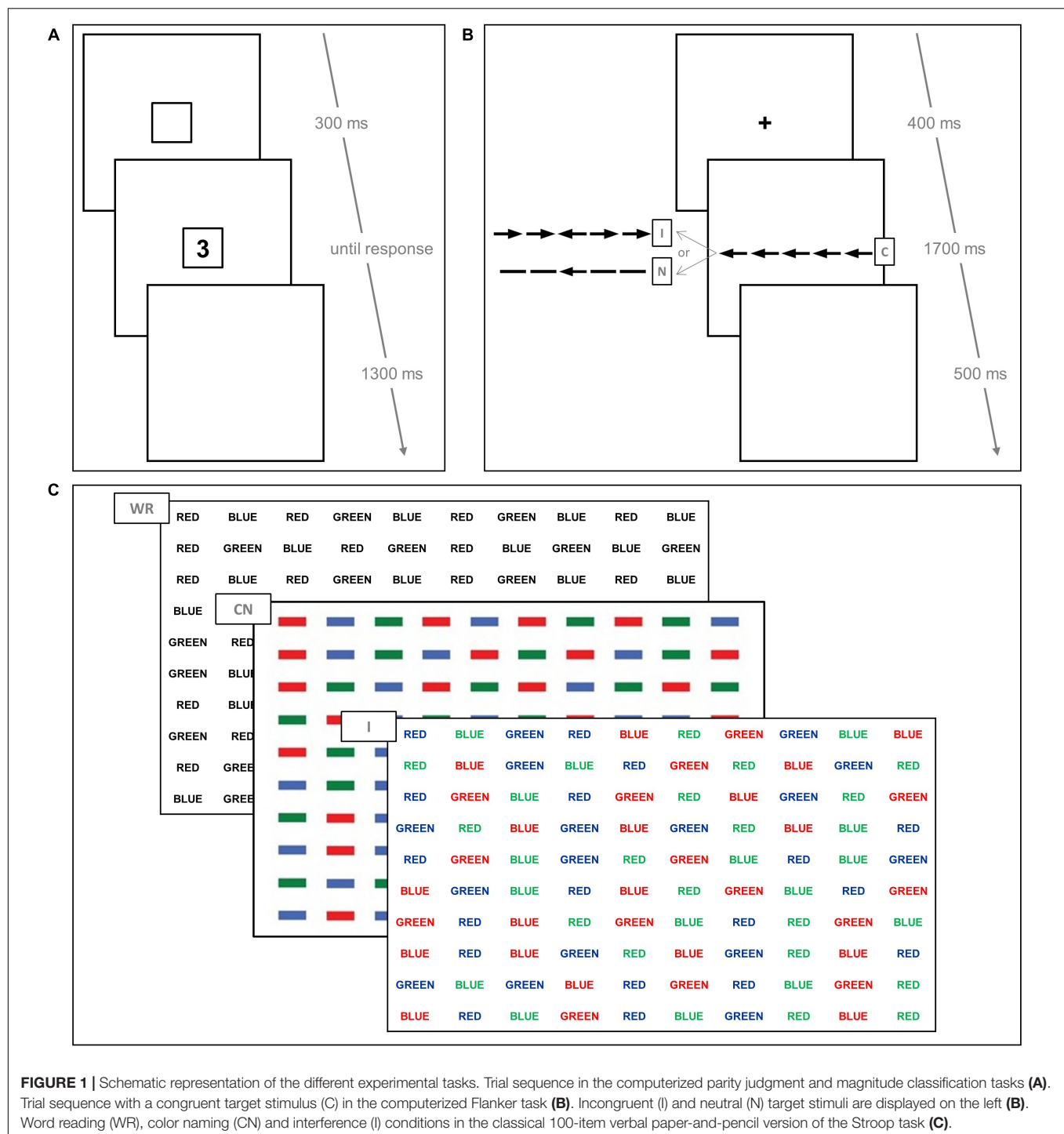
Stroop Task

The English adaptation of the classical 100-item verbal paper-and-pencil version of the Stroop paradigm was used to determine Stroop-like interference control (Stroop, 1935). The task consisted of three conditions, each comprising 100 items that were displayed in a 10×10 matrix on an A4 sheet of paper (see **Figure 1C**). In the word reading condition (WR), participants had to read color words (“red,” “blue,” “green”) printed in black ink. In the color naming condition (CN), they named swatches of red, blue and green ink. In the interference condition (I), participants were required to indicate the color of the ink (red, blue, green) that a color word (“red,” “blue,” “green”) was written in without reading the color word (e.g., they had to indicate “red” for the color word “green” printed in red ink). Participants were instructed to name/read the different items in each condition as quickly and as accurately as possible going from left-to-right. The time needed to complete each of the three conditions was recorded in every participant using a stopwatch. The WR and CN conditions served as control conditions.

To get a single inhibitory control measure indexing each participant’s Stroop effect, we calculated RT differences between the interference and color naming conditions. This is one of the standard methods for quantifying Stroop interference control

¹Since previous research indicated that split-half reliability was significantly lower for the parity than the magnitude SNARC effect (Georges et al., 2017), we decided to double the number of trials in the parity judgment task, as this was suggested to considerably enhance reliability estimates (Cipora and Wood, 2012; see also

Cipora and Nuerk, 2013; Cipora et al., 2016). Due to time constraints, we did, however, not increase the length of the magnitude classification task.



(Lansbergen et al., 2007). A greater RT difference is indicative of weaker interference control, as it reflects considerably slower RT in the interference than the color naming condition.

Flanker Task

The experiment was adapted from Eriksen and Eriksen (1974) and consisted of 48 trials (see **Figure 1B**). Each trial started with the display of a fixation cross (color: black; font: Arial;

point size: 28) in the center of a white screen. After 400 ms, a horizontal black arrow (height: 0.69°; width: 2.06°) was presented on a white background until response or for a maximum of 1700 ms. On half of the trials, the central arrow pointed in the left direction, while on the remaining half its pointing direction was reversed. Two black horizontal flanker arrows appeared on each side of the central arrow and pointed either in the same direction than the central arrow (i.e., congruent

condition, 16 trials) or in its opposite direction (i.e., incongruent condition, 16 trials). On the remaining neutral trials, the central arrow was flanked on both sides by two horizontal black bars. Participants were required to press the “A”/“L” key on a standard QWERTZ (Swiss-French) keyboard if the central arrow pointed in the left/right direction respectively. They were instructed to ignore the flanker arrows and bars. The inter-trial interval consisted of a blank screen of 500 ms. Trial sequence was identical for all participants and pseudo-randomized in a way that the correct response could not be the same more than 3 times consecutively. Moreover, the same target-distractor array did never successively appear. The actual experiment was preceded by 12 practice trials, consisting of 4 congruent, incongruent and neutral trials respectively. For each participant and every congruency condition, we computed error rates in percentages and averaged correct RTs that fell within 2.5 *SD* from the individual mean correct RT.

To incorporate error rates and RTs into a single performance measure, we computed inverse efficiency scores (IES) by dividing the means of congruent, incongruent or neutral correct RTs by their corresponding percentage accuracies for each participant (Bruyer and Brysbaert, 2011; Khng and Lee, 2014). IES thus adjusts RT performance for sacrifices in accuracy made in favor of response speed. Considering that faster responses together with fewer errors yield smaller IES, the smaller the IES is, the better the performance is.

To get a single inhibitory control measure indexing each participant's Flanker effect, we calculated individual IES differences by subtracting congruent from incongruent IES. A greater IES difference is indicative of weaker inhibitory control, as it reflects considerably worse performance (i.e., slower RT and/or more errors) in the incongruent compared to the congruent condition.

Speeded Matching-to-Sample Task

The speeded matching-to-sample task was used to determine general processing speed (GPS) and described in detail by Hoffmann et al. (2014a; see also Georges et al., 2016). Each trial consisted of a centrally displayed target shape and two possible solution shapes, displayed below to the left and right. Participants had to identify the solution that was identical to the target as quickly as possible by clicking the “A”/“L” key on a QWERTZ keyboard if it appeared on the bottom left/right respectively. For each participant, we averaged correct RTs that fell within 2.5 *SD* from the individual mean correct RT.

RESULTS

Descriptives

SNARC Effects

Split-half reliabilities were calculated for the parity and magnitude SNARC effect regression slopes using the odd-even method to control for systematic influences of practice or tiring within the tasks (see Cipora and Nuerk, 2013; Cipora et al., 2016; Georges et al., 2016, 2017; Ninaus et al.,

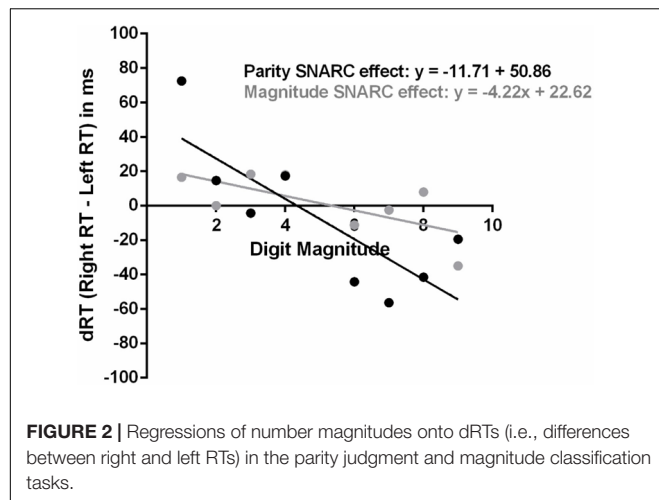
2017). Trials were odd-even half-split (based on order of appearance) and two SNARC effect regression slopes were calculated separately for each participant and each task. The correlation coefficients were Spearman-Brown corrected to get a reliability estimate for the entire set of items. The Spearman-Brown corrected correlation coefficient was $r = 0.56$ in both the parity judgment and magnitude classification tasks.

To determine whether relatively low reliabilities could be caused by the influence of bivariate outliers, we performed linear regression analyses between odd and even SNARC effect regression slopes and identified influential data points based on the conventional Cook's distances criterion of $>4/N$ (Cook, 1979; Bollen and Jackman, 1985; see Viarouge et al., 2014 for application of this method in the SNARC context). Two separate analyses were performed – one for the parity judgment task and one for the magnitude classification task. For the parity judgment task, analysis revealed two influential data points with Cook's distances greater than .154 (i.e., $4/26$). After removal of these participants, the bivariate correlation between odd and even parity SNARC effect regression slopes remained similar ($r = 0.35$ for $N = 24$; $r = 0.39$ for $N = 26$; Fisher's z for comparison of two correlations based on independent groups: $z = 0.15$; $p = 0.88$), yielding a Spearman-Brown corrected reliability estimate of $r = 0.52$. For the magnitude classification task, three influential cases were identified with Cook's distances greater than .154 (i.e., $4/26$). After removal of these three influential data points, the correlation between odd and even magnitude SNARC effect regression slopes improved from $r = 0.39$ (for $N = 26$) to $r = 0.53$ (for $N = 23$), yielding a Spearman-Brown corrected reliability estimate of $r = 0.7$. Influential cases were not removed in any of the following correlation analyses, where $N = 26$.

The mean SNARC effect regression slope across all participants was significantly negative in the parity judgment but not the magnitude classification task [parity SNARC effect regression slope = -11.71 ; $SD = 13.36$; $t(25) = -4.47$; $p < 0.001$; magnitude SNARC effect regression slope = -4.22 ; $SD = 12.89$; $t(25) = -1.67$; $p = 0.11$; see **Figure 2**]. A repeated-measures ANOVA on the SNARC effect regression slopes also revealed a main effect of task [$F(1,25) = 4.59$; $p = 0.042$; $\eta_p^2 = 0.16$], indicating stronger number-space associations in the parity judgment than the magnitude classification task in terms of the inclination of the regression lines. Overall, a large proportion of the participants displayed a negative SNARC effect regression slope in both the parity judgment (20/26; 76.92%) and magnitude classification tasks (18/26; 69.23%).

Stroop Effect

The mean RTs across all participants were 43.04 s ($SD = 8.19$) in the word reading, 64.19 s ($SD = 10.08$) in the color naming and 97.04 s ($SD = 18.96$) in the interference conditions. A repeated measures ANOVA on RT including condition as within-subject variable revealed a main effect of condition [$F(2,50) = 222.93$; $p < 0.001$; $\eta_p^2 = 0.9$]. Participants performed significantly worse in the interference compared to the color naming [$t(25) = -13.56$; $p < 0.001$] and the word reading



[$t(25) = -16.13$; $p < 0.001$] conditions. Performances were also significantly lower during color naming than word reading [$t(25) = -12.51$; $p < 0.001$].

The mean RT difference between the interference and color naming conditions (i.e., Stroop effect) across all participants was 32.85 s ($SD = 12.35$). Individual Stroop effects were used for the subsequent correlation analyses.

Flanker Effect

As for the two SNARC effects, reliability of the Flanker effect was determined using the split-half method (Greene et al., 2008; see also MacLeod et al., 2010). More concretely, congruent and incongruent trials were odd-even half-split (based on order of appearance) and Flanker effects (i.e., differences between incongruent and congruent IES) were computed separately for each half in every participant. The correlation between IES differences (i.e., Flanker effects) calculated on odd and even trials was Spearman-Brown corrected, yielding a reliability estimate of $r = 0.63$.

The mean error rates and RTs across all trials and participants were 1.2% ($SD = 3.07$) and 432 ms ($SD = 72$) in the congruent, 8.41% ($SD = 11.17$) and 490 ms ($SD = 67$) in the incongruent and 0.72% ($SD = 2.04$) and 440 ms ($SD = 64$) in the neutral conditions respectively. Error rates and RTs did not correlate in the congruent ($r = 0.01$; $p = 0.96$) and neutral ($r = -0.23$; $p = 0.25$) conditions, suggesting that these performance estimates provide different aspects of inhibitory control. Moreover, there was a speed-accuracy trade-off in the incongruent condition ($r = -0.53$; $p = 0.006$).

A repeated measures ANOVA on IES including congruency condition as within-subject variable revealed a main effect [$F(2,50) = 47.00$; $p < 0.001$; $\eta_p^2 = 0.65$]. Participants performed significantly worse on incongruent (IES = 538.69 ms; $SD = 67.57$) compared to congruent [IES = 438.09 ms; $SD = 75.49$; $t(25) = -7.33$; $p < 0.001$] and neutral [IES = 443.17 ms; $SD = 62.69$; $t(25) = 7.13$; $p < 0.001$] trials. Performances did not differ between the congruent and neutral conditions.

The mean IES difference between incongruent and congruent trials (i.e., Flanker effect) across all participants was 100.59 ms

($SD = 69.99$). Individual Flanker effects were used for the subsequent correlation analyses.

General Processing Speed

The mean RT across all trials and participants in the speeded matching-to-sample task was 626 ms ($SD = 242$). RTs significantly positively correlated with RTs on the parity judgment (613 ms; $SD = 85$; $r = 0.7$; $p < 0.001$), magnitude classification (536 ms; $SD = 74$; $r = 0.46$; $p = 0.019$), Stroop (68.09 s; $SD = 11.11$; $r = 0.4$; $p = 0.045$), and Flanker tasks (454 ms; $SD = 66$; $r = 0.41$; $p = 0.036$). It thus provided a valid index of general processing speed and can be used as a control measure in a partial correlation analysis to verify whether potentially significant correlations between number-space associations and any of the interference control measures might be reduced to inter-individual differences in general processing speed.

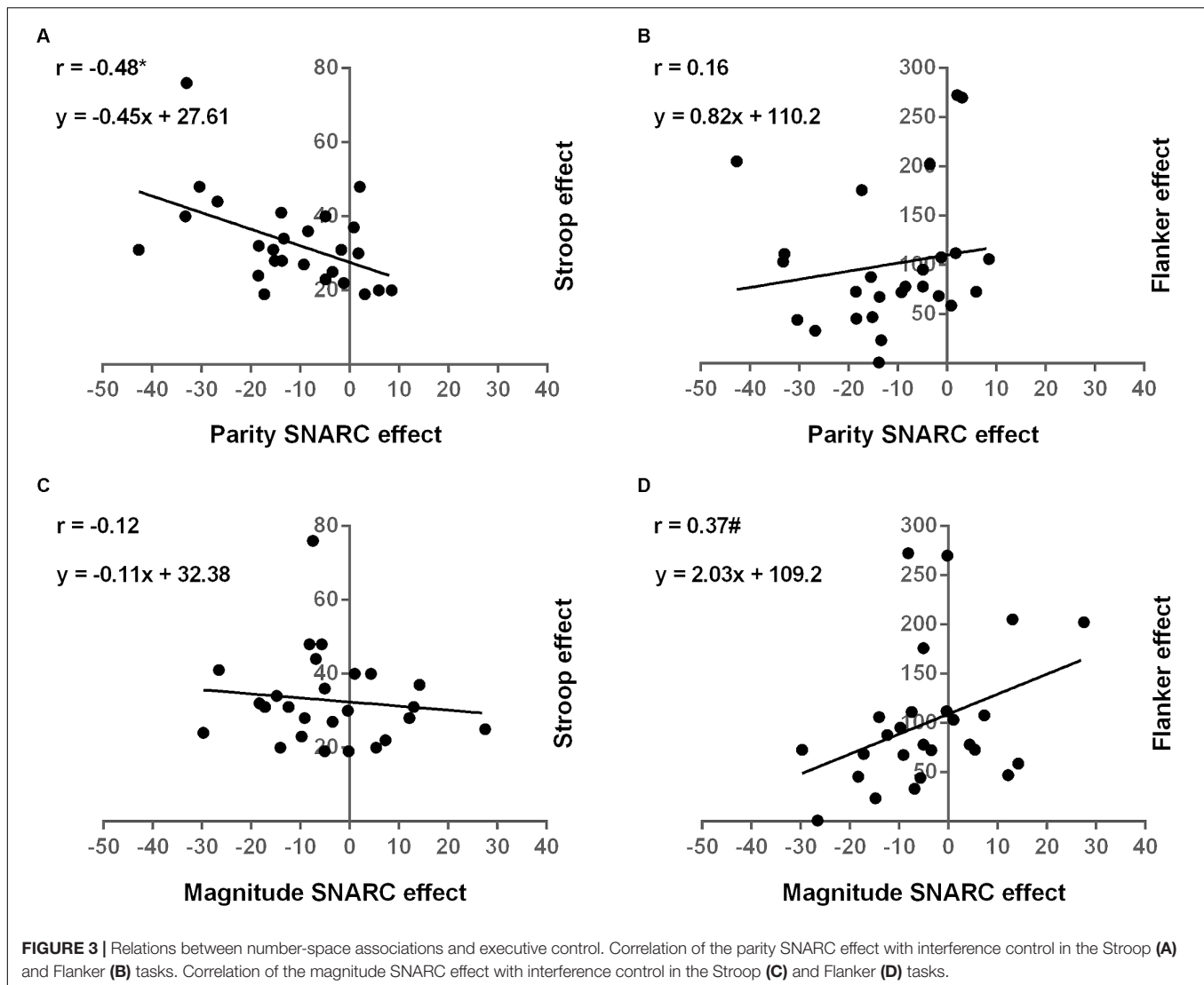
All descriptive information is displayed in **Table 1**.

Correlation Analyses

All reported correlations are two-tailed, unless otherwise stated. Stronger parity SNARC effects were associated with weaker interference control in the Stroop task ($r = -0.48$; $p = 0.012$; **Figure 3A**). Conversely, no relation was observed between the parity SNARC effect and interference control in the Flanker task ($r = 0.16$; $p = 0.44$; **Figure 3B**). This difference between the relations of the parity SNARC effect with interference control in the Stroop and Flanker paradigms reached significance, as revealed by Pearson and Filon's z (Pearson and Filon, 1898), assessing differences between two overlapping correlations based on dependent samples ($z = -2.51$; $p = 0.006$; one-tailed). As opposed to number-space associations in implicit tasks, stronger magnitude SNARC effects trended to be associated with better interference control in the Flanker task ($r = 0.37$; $p = 0.06$; **Figure 3D**). The magnitude SNARC effect was, however, unrelated to interference control in the Stroop task ($r = -0.12$; $p = 0.58$; **Figure 3C**). The difference between the correlations of the magnitude SNARC effect with Stroop and Flanker effects was significant ($z = -1.80$; $p = 0.04$; one-tailed). In line with previous findings, the parity and magnitude SNARC effects did not correlate ($r = 0.08$; $p = 0.7$). The difference between the relations of the Stroop effect with the parity and magnitude SNARC effects, however, only trended toward significance ($z = -1.54$; $p = 0.06$; one-tailed). Likewise, no significant difference could be observed between the correlations of the SNARC effects in implicit and

TABLE 1 | Descriptive information.

Variable	All participants
Gender (f/m)	15/11
Age (years)	26.86 ($SD = 3.29$; range = 22.17 – 33.43)
Parity SNARC effect	-11.71 ($SD = 13.36$; range = -42.70 – 8.49)
Magnitude SNARC effect	-4.22 ($SD = 12.89$; range = -29.70 – 27.50)
Stroop effect (s)	32.85 ($SD = 12.35$; range = 19 – 76)
Flanker effect (ms)	100.59 ($SD = 69.99$; range = 1 – 272.56)
General processing speed (ms)	626 ($SD = 242$; range = 381 – 1372)



explicit tasks with the Flanker effect ($z = -0.87$; $p = 0.19$; one-tailed). Performances on the Stroop and Flanker tasks did also not correlate ($r = -0.14$; $p = 0.5$), confirming qualitative differences between these interference measures. Finally, general processing speed did not relate to any of the SNARC effects or inhibitory control measures (all $ps > 0.05$).

Considering the non-perfect reliabilities of the SNARC effect regression slopes, we corrected bivariate correlations for attenuation using Spearman's correction for attenuation formula, corresponding to $r_{xy} / \sqrt{r_{xx} * r_{yy}}$, with r_{xx} and r_{yy} coding for the reliabilities of X and Y respectively (Spearman, 1904, 1910; Muchinsky, 1996; see also Cipora and Nuerk, 2013; Gloria et al., 2016; Georges et al., 2017, for a comparable application of this correction for attenuation method). This procedure determines the correlation between two variables if they were perfectly reliable, and therefore provides for a more accurate estimate of the correlation between two parameters. Attenuated and disattenuated correlation coefficients are shown in the upper and lower part of Table 2 respectively.

All the above relations remained similar when controlling for general processing speed in a partial correlation analysis (see Table 3).

DISCUSSION

Inter-Individual Differences in Number-Space Associations During Parity Judgments Relate to Interference Control in the Stroop Task

Stronger number-space associations in the parity judgment task correlated with weaker interference control in the Stroop task in young adults with diagnosed or self-reported ADHD. This relation remained significant even after controlling for general processing speed, previously implicated in both the parity SNARC (e.g., Wood et al., 2008; Cipora and Nuerk, 2013; Hoffmann et al., 2014b) and Stroop effects (e.g., Bugg et al.,

TABLE 2 | Correlation analysis.

	1	2	3	4	5
(1) Parity SNARC effect	–	0.08	–0.48*	0.16	–0.12
(2) Magnitude SNARC effect	0.14	–	–0.12	0.37#	–0.28
(3) Stroop effect	–0.64	–0.16	–	–0.14	0.04
(4) Flanker effect	0.27	0.62	–0.18	–	–0.22
(5) General processing speed	–0.16	–0.37		–0.28	–

Attenuated correlation coefficients are displayed in bold in the upper part of the table. Disattenuated correlation coefficients are displayed in the lower part of the table. * $p < 0.05$; # $p = 0.06$.

TABLE 3 | Partial correlation analysis controlling for general processing speed.

	1	2	3	4
(1) Parity SNARC effect	–	0.05	–0.48*	0.14
(2) Magnitude SNARC effect		–	–0.11	0.33
(3) Stroop effect			–	–0.13
(4) Flanker effect				–

* $p < 0.05$.

2007; Hoffmann et al., 2014b). The present findings extend the recently reported relation between stronger parity SNARC effects and weaker Stroop inhibitory control in the elderly and confirm the hypothesis that the null relation in young healthy participants can be explained by their near ceiling performances on the Stroop task (Hoffmann et al., 2014b).

In contrast, number-space associations during parity judgments were not related to interference control in the Flanker task in the present population. It is unlikely that this null relation can be explained by insufficient variance in the Flanker effect due to near ceiling task performances, considering the tendency for a positive relation between interference control in the Flanker task and number-space associations in the magnitude classification task (discussed in the next section). Moreover, individuals with ADHD were previously shown to feature abnormal inhibitory control in both the Stroop (Nigg et al., 2005; Walker et al., 2000; King et al., 2007) and Flanker paradigms (Lundervold et al., 2011). The spatial coding mechanisms underlying the parity SNARC effect thus depend on those inhibitory control processes indexed by the Stroop but not the Flanker effect. Overall, this provides valuable information regarding the type of conflict encountered during parity judgments, thereby advancing our understanding of the spatial coding processes underlying the parity SNARC effect.

To characterize the coding mechanisms accounting for the parity SNARC effect, it is important to firstly understand the cognitive processes underlying interference control in the Stroop and Flanker tasks. Interference in the Stroop paradigm originates at the semantic level from an attribute that is intrinsic to the target stimulus (i.e., the meaning of the color word conflicts with the semantic representation of the ink color, e.g., Klein, 1964; La Heij, 1988). Moreover, the distracting color word meaning is highly salient, considering that literate individuals are primed to automatically access a word's meaning upon sight prior to processing any additional features (Ashcraft

and Radvansky, 2010). Conversely, interference in the Flanker paradigm occurs spatially instead of semantically from lateral arrows that are drawn from the same set of stimuli than the target stimulus (Eriksen and Schultz, 1979). The relation between the parity SNARC and Stroop (but not Flanker) effects thus suggests that the spatial code associated with numerical magnitude during parity judgments is semantic in nature and/or intrinsic to the target stimulus (see Table 4). Since the Stroop as opposed to the Flanker task yields basically no perceptual interference (Valle-Inclán, 1996), the conflict in the parity judgment task is also unlikely of perceptual nature. This outcome is in line with the parity judgment paradigm, where the task-relevant parity status and the conflicting spatial code associated with the automatically activated yet task-irrelevant magnitude information reflect distinct semantic properties of the same target number.

While distraction in the Flanker task is provided by externally available visuospatial information (i.e., the flanking arrows), the distracting color word meaning in the Stroop paradigm is rather verbal in nature. The Stroop task is highly left lateralized, most prominently in the left dorsolateral prefrontal cortex and inferior frontal areas, previously implicated in the resolution of verbal conflict (Jonides et al., 1998; Leung et al., 2000; Jonides and Nee, 2006). In the present Stroop paradigm, responses were also given verbally, thereby adding to the already rather verbal nature of the Stroop task. The strong relation between the parity SNARC and Stroop effects thus suggests that the distracting spatial code associated with numerical magnitude in the parity judgment task might also be verbal in nature (see Table 4). In line with previous claims, this suggests that the parity SNARC effect predominantly results from verbal-spatial polarity coding as opposed to arising from the spatial coding of numerical magnitudes on a horizontally oriented MNL (Gevers et al., 2010; Georges et al., 2017).

According to the dimensional overlap model by Kornblum et al. (1990; see also Kornblum and Lee, 1995; Zhang et al., 1999), interference in the Flanker task mainly reflects a stimulus-stimulus conflict, where the pointing directions of the task-irrelevant flanking arrows interfere with that of the targeted central arrow at the early stage of stimulus encoding. Such interference is likely resolved via the spatial filtering of the perceptual distractors and the narrowing of the attentional focus to the task-relevant central arrow location (Wendt et al., 2012). Conversely, conflict in the Stroop paradigm occurs at multiple stages of stimulus processing (Zhang and Kornblum, 1998; Milham et al., 2001; De Houwer, 2003). In addition to the semantic stimulus-stimulus conflict at earlier processing stages (e.g., Klein, 1964; Kornblum et al., 1990; Sharma and McKenna, 1998; Schmidt and Cheesman, 2005; Goldfarb and Henik, 2007), stimulus-response conflict arises during response selection (e.g., Cohen et al., 1990; MacLeod, 1991; van Veen and Carter, 2005; Szucs and Soltész, 2010), when the task-relevant ink color and the irrelevant meaning of the color word activate competing responses. Such stimulus-response conflict is then probably resolved via biasing units reflecting the task-relevant semantic dimension (i.e., the ink color of the color word; Szucs et al., 2009). The relation between number-space associations

in the parity judgment task and interference control in the Stroop paradigm thus suggests that the parity SNARC effect also mainly originates at later processing stages during response selection (see **Table 4**). Accordingly, the response provoked by the task-irrelevant numerical magnitude-associated spatial code competes/conflicts with that induced by the task-relevant parity status prior to response execution. Such competition is likely resolved via biasing units coding the response associated with the task-relevant parity status (see **Table 4**). Considering the absence of a relation between the parity SNARC and Flanker effects, interference in the parity judgment task is unlikely controlled by filtering mechanisms already at the early stage of number encoding. This outcome is in line with previous models proposed to account for the parity SNARC effect (Keus et al., 2005; Gevers et al., 2006). According to Gevers et al. (2006), the parity SNARC effect results from the interference of two processing routes operating in parallel. The conditional route links task-relevant parity information with response keys based on task instructions, while the unconditional route conveys the automatic association between numerical magnitude and space. On congruent trials, both routes activate the same response location, while on incongruent trials responses are slowed down and more error-prone since the two routes activate competing outcomes.

Evidence for such parallel processing of task-relevant and irrelevant information and of conflict resolution mainly at the response selection stage during parity judgments has also been provided by EEG studies. Namely, congruency effects were previously reported on the latency of the lateralized readiness potential (Keus et al., 2005; Gevers et al., 2006), an EEG component considered to be the output of response selection stages (Gratton et al., 1988; Coles, 1989; for a review, see also Leuthold et al., 2004). In addition and in line with observations regarding the Stroop effect (Ilán and Polich, 1999; Zurrón et al., 2009; for a review, see Sahinoglu and Dogan, 2016), the P300 peak latency did not show an onset difference between congruent and incongruent trials in the parity judgment task (Gevers et al., 2006), indicating that the conflict indexed by the parity SNARC effect is unlikely detected at early perceptual stages.

The assumption that the conflict indexed by the parity SNARC effect originates at later processing stages during response selection also agrees with findings regarding stronger parity SNARC effects in the elderly compared to young healthy individuals (Hoffmann et al., 2014b; Ninaus et al., 2017). Elderly persons featured weaker interference control in the Stroop paradigm than young controls (West and Alain, 2000; Van der Elst et al., 2006), suggesting an age-associated decline in conflict resolution particularly at later response selection stages. In contrast, the resolution of stimulus-stimulus conflict at earlier processing stages in the Flanker task did not differ between younger and older participants as reflected by similar behavioral performances of both age groups (Falkenstein et al., 2001; Nieuwenhuis et al., 2002).

Inter-Individual Differences in Number-Space Associations During Magnitude Classifications Relate to Interference Control in the Flanker Task

Inter-individual variability in the strength of number-space associations during explicit magnitude classifications did not relate to inter-individual differences in the Stroop effect. Conversely, stronger magnitude SNARC effects were associated with better interference control in the Flanker task. However, it should be noted that this correlation did not reach significance, also not prior to partialling out the effects of general processing speed. Nonetheless, the relation between more pronounced number-space associations during explicit magnitude classifications and better interference control in the Flanker paradigm was significantly different from the null correlation between the magnitude SNARC and Stroop effects.

The latter null relation might suggest that the spatial code associated with numerical magnitude during explicit classifications is not of verbal nature, akin to the verbal interference encountered in the Stroop paradigm (Jonides et al., 1998; Leung et al., 2000; Jonides and Nee, 2006) and probably also during parity judgments. This lines up with previous findings indicating that the magnitude SNARC effect

TABLE 4 | Characteristics of the spatial code associated with numerical magnitude during parity judgments and magnitude classifications.

Spatial code characteristics	Parity judgment		Magnitude classification	
	Relevance	Irrelevant	Alternative a	
	Nature	Verbal	Alternative b	
	Origin	Intrinsic to target number		
	Processing stage	Response selection		
	Processing mechanism	Suppression via biasing units reflecting relevant response	Suppression via spatial filtering	Activation via selective attention

In the parity judgment task, the irrelevant verbal-spatial code associated with the numerical magnitude of the target number interferes with the spatial location of the response based on parity status during response selection. Interference is resolved via biasing units coding the spatial location of the relevant response. For the magnitude classification task, two alternatives (a and b) are outlined. According to alternative (a), the irrelevant visuospatial codes associated with the numerical magnitudes of the numbers represented adjacently to the target number on the MNL interfere with the processing of the target number during encoding. Interference is resolved via the spatial filtering of the irrelevant numerical magnitude representations on the MNL. According to alternative (b), the relevant visuospatial code associated with the numerical magnitude of the target number is activated during encoding via selective attention.

was selectively abolished by a visuospatial but not verbal WM load, highlighting the importance of visuospatial coding mechanisms (van Dijck et al., 2009). Moreover, Georges et al. (2017) reported a relation between stronger magnitude SNARC effects and greater preferences for spatial as opposed to object visualization. Number-space associations during explicit magnitude classifications thus likely predominantly depend on visuospatial processing resources in the right parietal cortex associated with spatial visualization (Lamm et al., 1999; see **Table 4**). The absence of a correlation between number-space associations in the magnitude classification task and interference control in the Stroop paradigm might also indicate that the magnitude SNARC effect differs from conflict that originates from a semantic feature intrinsic to the target stimulus (i.e., the central number). Furthermore, interference in the magnitude classification task might diverge from conflict that is mainly resolved at the response selection stage, such as the conflict induced by the irrelevant color word meaning in the Stroop paradigm. The null relation between the magnitude SNARC and Stroop effects could, however, also simply suggest that no conflict arises from the spatial code associated with numerical magnitude during explicit classifications.

When considering the tendency for an association between the magnitude SNARC and Flanker effects, it might suggest that the potential interference during explicit classifications originates from irrelevant visuospatial information extrinsic to the target stimulus (see **Table 4**, alternative a). Additionally, it could indicate conflict resolution directly at the early stage of stimulus encoding via spatial filtering (see **Table 4**, alternative a). At first, this idea seems difficult to reconcile with the magnitude classification paradigm, considering that it only comprises a single task-relevant centrally displayed number. If extrinsic distraction might be encountered during magnitude classifications, it can only originate internally. One possibility is for instance that interference arises from task-irrelevant numerical magnitudes represented adjacently to the target number on a horizontal MNL (or sequence within WM; see e.g., Fias et al., 2011). Indirect support for such an interplay between the externally available task-relevant number and internally represented task-irrelevant numerical magnitudes was provided by Nuerk et al. (2005). Their findings suggested that the representation of closely related task-irrelevant numbers can interfere with task-relevant numerical magnitude classifications at least when these distracting numbers are externally available. Of course, the assumption of such interference by internally represented task-irrelevant numerical magnitudes is only valid if the spatial code associated with numerical magnitude during explicit classifications is indeed visual instead of verbal in nature. A greater ability to suppress such task-irrelevant spatial-numerical activations at earlier processing stages (akin to the spatial filtering of distractors in the Flanker task) might then facilitate the processing of the task-relevant numerical magnitude together with its associated spatial code, manifesting in stronger magnitude SNARC effects. This explanation could then account for the positive relation between stronger magnitude SNARC effects and better interference control in the Flanker task.

Alternatively, the trend for a relation between stronger magnitude SNARC effects and better inhibitory control in the Flanker task might indicate that a greater ability to selectively focus attention on task-relevant information (as indexed by better interference control in the Flanker task; see Wendt et al., 2012) is associated with stronger number-space associations during explicit magnitude classifications. Of course, this entails that the spatial code associated with the task-relevant numerical magnitude is also relevant rather than distracting for successful resolution of the magnitude classification task (see **Table 4**, alternative b). The relevance of spatial-numerical mappings during explicit magnitude classifications could then also account for the lack of a correlation between the magnitude SNARC and Stroop effects. Moreover, it seems likely considering that coding small/large numerical magnitudes as left/right on the MNL (or within WM) might assist left-/right-sided numerical magnitude classifications. It would also provide an explanation for the observation that stronger magnitude SNARC effects are not related to weaker arithmetic performances (Georges et al., 2017), contrary to the parity SNARC effect (e.g., Cipora et al., 2016; Georges et al., 2017; but see Cipora and Nuerk, 2013). In general, more linear spatial representations of numerical magnitudes, as assessed using number line estimations, are commonly associated with better magnitude comparison performances (Laski and Siegler, 2007) as well as higher math skills (Link et al., 2014). These findings thus highlight the importance/relevance of spatial-numerical representations for arithmetic performances.

Intra-Individual Differences in Number-Space Associations and Task-Dependent Differences in the Relation to Interference Control

The present results provide further evidence for the previously reported intra-individual variability in number-space associations depending on the implicit or explicit nature of numerical magnitude processing (van Dijck et al., 2009; Georges et al., 2017). More concretely, parity and magnitude SNARC effects were uncorrelated and related (or at least tended to relate) inversely to distinct inhibitory control measures, namely negatively with the Stroop and positively with the Flanker effects respectively. This heterogeneity in the cognitive processes underlying the SNARC effect generally agrees with studies indicating that both long-term spatial coding mechanisms such as the spatial representation of numerical magnitudes on a MNL and temporary associations between the ordinal position of numerical magnitudes and space in WM might exist in parallel (Ginsburg and Gevers, 2015; Huber et al., 2016; but see Abrahamse et al., 2016).

Previous explanations for such intra-individual variations in number-space associations depending on the number processing task suggested task-related differences in the nature of the numerical magnitude-associated spatial code, with verbal- and visuospatial coding processes probably underlying the parity and magnitude SNARC effects respectively (van Dijck et al., 2009; Gevers et al., 2010; Georges et al., 2017). This assumption might further be supported by the present findings. Namely,

only the parity SNARC effect correlated with interference control in the Stroop paradigm, reflecting the suppression of task-irrelevant verbal information (i.e., the color word meaning).

The current results, however, allow for an additional (or even alternative) explanation regarding intra-individual differences in number-space associations depending on the number processing task. Namely, as already discussed above, the relations of the parity and magnitude SNARC effects with stimulus–response and stimulus–stimulus conflict resolution in the Stroop and Flanker paradigms respectively suggests that the task-dependency of number-space associations might result from task-related differences in the processing stages of the spatial code associated with numerical magnitude, irrespective of its visual or verbal nature. While the conflict provided by the numerical magnitude-associated spatial code during parity judgments might predominantly be resolved at the response selection stage via biasing units coding the task-relevant response location (see **Table 4**), the potential conflict during explicit magnitude classifications probably rather originates from extrinsic distractors and is resolved via their spatial filtering at earlier processing stages (see **Table 4**, alternative a). The conflicts indexed by the parity and magnitude SNARC effects would thus have distinct origins and be resolved via different mechanisms at different processing stages, thereby potentially explaining the task-dependency of number-space associations.

Alternatively, as already mentioned before, differences in the relevance of the spatial code associated with numerical magnitude during parity judgments and magnitude classifications and consequently in its processing (inhibition vs. activation respectively) could probably underlie the task-dependency of number-space associations (see **Table 4**).

Limitations and Future Directions

First, it should be reminded that split-half reliabilities for both the parity and magnitude SNARC effect regression slopes were relatively low. Lower reliabilities are, however, not unusual in SNARC-related studies. Comparably low reliabilities were also reported in previous studies by means of both internal consistency (Cipora and Nuerk, 2013; Viarouge et al., 2014; Georges et al., 2016, 2017; Cipora et al., 2018) as well as test–retest stability (Viarouge et al., 2014).

To increase reliability estimates, the length of the parity judgment task was increased, which was shown to yield better split-half reliability estimates (Cipora and Wood, 2012, 2017; see also Cipora and Nuerk, 2013; Cipora et al., 2016). Nonetheless, the Spearman–Brown corrected correlation coefficient in the present parity judgment task was comparable to that in Georges et al. (2017) using a task that included only half of the number of trials. Moreover, similar split-half reliability estimates were obtained for the parity and magnitude SNARC effects, albeit the parity judgment task had twice the length of the magnitude classification paradigm. Increasing the number of repetitions per stimulus in the parity judgment task did thus not seem to enhance split-half reliability in the current study. It should, however, be noted that the present study only included individuals with diagnosed or self-reported ADHD, generally featuring relatively

high intra-individual variability in RTs (Castellanos et al., 2005; Vaurio et al., 2009). This might thus have generally accounted for the lower reliabilities, despite the increase in some of the task lengths.

Importantly, the relatively poor reliabilities of the parity and magnitude SNARC effect regression slopes could have negatively impacted the correlations reported in the current study. Namely, the upper bound of a correlation between two parameters depends on their reliabilities in that the highest correlation between two variables equals the square root of the product of their reliabilities [i.e., $\sqrt{r_{xx} \cdot r_{yy}}$, with r_{xx} and r_{yy} coding for the reliabilities of X and Y respectively]. The correlation between two variables is thus weakened by measurement error, such that true correlations between measures with poor reliability might be overlooked (Osborne and Waters, 2002, p. 2). Consequently, we need to be careful when drawing conclusions about (the absence of) relations between number-space associations and interference control from the present findings. Nevertheless, any task-related differences in the relations between number-space associations and the different interference control measures cannot be ascribed to low measurement reliability, since split-half reliability estimates for the parity and magnitude SNARC effect regression slopes were equally low.

Another drawback of the present study could be the relatively small sample size of $N = 26$. A *post hoc* power analysis based on effect size, conventional alpha level, and sample size (i.e., $N = 26$) using the program G*Power (Faul et al., 2007, 2009) revealed that the probability of rejecting a false null hypothesis was 81% for large ($r = 0.5$), 34% for medium ($r = 0.3$) and 8% for small ($r = 0.1$) effect sizes. The present study had thus sufficient power to detect a significant relation between the SNARC effect and inhibitory control at the large effect size level. Conversely, less than adequate statistical power was obtained at the small to medium effect size level to reject an incorrect null hypothesis. The lack of sufficient power for detecting small to medium effect sizes could potentially account for the non-significant relation between stronger number-space associations during magnitude classifications and better interference control in the Flanker task in the current sample.

Future studies should also consider the inclusion of control variables. Especially the involvement of verbal and visuospatial WM could be assessed in greater detail. Relations between number-space associations and inhibitory control might indeed be (partially) confounded by WM processes. WM is not only implicated in the Stroop (Long and Prat, 2002; Kane and Engle, 2003; Hutchison, 2011) as well as Flanker effects (Redick and Engle, 2006; Heitz and Engle, 2007), but also likely contributes to number-space associations (van Dijck et al., 2009). Nonetheless, it should be noted that Hoffmann et al. (2014b) controlled for the influence of verbal WM in their study, thereby excluding the possibility that the relation between stronger parity SNARC effects and weaker interference control in the Stroop paradigm might be confounded by verbal WM.

Future research could also elaborate on the assumption that no interference originates from the spatial code associated with numerical magnitude during explicit classifications by

investigating whether the strength of the magnitude SNARC effect varies with age, similarly to the age-associated increase in number-space associations during parity judgments (Hoffmann et al., 2014b; Ninaus et al., 2017). Inhibitory control declines with age (see Glisky, 2007) mostly regarding conflict resolution at later response selection stages (Falkenstein et al., 2001; Nieuwenhuis et al., 2002; Van der Elst et al., 2006), while target selection processes usually remain intact even in the elderly (West and Alain, 2000). Consequently, if the magnitude SNARC effect indeed does not index interference control, its strength should not be altered by aging.

CONCLUSION

Stronger parity SNARC effects were associated with weaker interference control in the Stroop but not Flanker task in young adults with diagnosed or self-reported ADHD. Number-space associations in the parity judgment task thus index conflict resolution akin to the Stroop effect. In other terms, the parity SNARC effect likely reflects interference between the (probably) verbal-spatial code associated with numerical magnitude and the spatial location of the response associated with parity status at later processing stages during response selection (see **Table 4**). Conversely, the magnitude SNARC effect was not related to interference control in the Stroop paradigm. Stronger number-space associations during explicit magnitude classifications, however, tended to be associated with better conflict resolution in the Flanker task. The (probably) visuospatial code associated with numerical magnitude is thus likely relevant during explicit magnitude classifications, with its activation at the early stage of stimulus encoding underlying the magnitude SNARC effect (see **Table 4**, alternative b). Overall, the present findings suggest that the relevance/importance of

number-space associations for numerical judgments depends on the implicit or explicit nature of the number processing task. While the spatial code associated with numerical magnitude seems to assist explicit magnitude classifications (and is therefore activated at the encoding stage), it seems to interfere with parity judgments (and is therefore suppressed at the response selection stage). Such differences in the relevance of the numerical magnitude-associated spatial code during parity judgments and magnitude classifications and in the related executive control mechanisms monitoring its processing (suppression vs. activation respectively) might account for the previously reported task-dependency of number-space associations.

AUTHOR CONTRIBUTIONS

CG, DH, and CS: conceived and designed the experiments. CG: analyzed the data. CG, DH, and CS: wrote the paper.

FUNDING

The current research was supported by the National Research Fund Luxembourg (FNR; www.fnr.lu) under Grant AFR PHD-2012-2/4641711.

ACKNOWLEDGMENTS

Some content in the introduction and method sections has been taken and/or adapted from Georges' dissertation thesis, defended on 17th February 2017 in Luxembourg. A full reference to the thesis is provided in the reference list.

REFERENCES

- Abrahamse, E., van Dijk, J.-P., and Fias, W. (2016). How does working memory enable number-induced spatial biases? *Front. Psychol.* 7:977. doi: 10.3389/fpsyg.2016.00977
- Adler, L. A., Spencer, T., Faraone, S. V., Kessler, R. C., Howes, M. J., Biederman, J., et al. (2006). Validity of pilot adult ADHD self-report scale (ASRS) to rate adult ADHD symptoms. *Ann. Clin. Psychiatry* 18, 145–148. doi: 10.1080/10401230600801077
- Ashcraft, M. H., and Radvansky, G. A. (2010). *Cognition*, 5th Edn. Boston, MA: Prentice Hall.
- Bollen, K. A., and Jackman, R. W. (1985). Regression diagnostics: an expository treatment of outliers and influential cases. *Sociol. Methods Res.* 13, 510–542. doi: 10.1177/0049124185013004004
- Bruyer, R., and Brysbaert, M. (2011). Combining speed and accuracy in cognitive psychology: is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)? *Psychol. Belg.* 51, 5–13. doi: 10.5334/pb-51-1-5
- Bugg, J. M., DeLosh, E. L., Davalos, D. B., and Davis, H. P. (2007). Age differences in Stroop interference: contributions of general slowing and task-specific deficits. *Aging Neuropsychol. Cogn.* 14, 155–167. doi: 10.1080/138255891007065
- Bull, R., Marschark, M., and Blatto-Valle, G. (2005). SNARC hunting: examining number representation in deaf students. *Learn. Individ. Dif.* 15, 223–236. doi: 10.1016/j.lindif.2005.01.004
- Carlson, S. M., and Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Dev.* 72, 1032–1053. doi: 10.1111/1467-8624.00333
- Castellanos, F. X., Sonuga-Barke, E. J. S., Scheres, A., Di Martino, A., Hyde, C., and Walters, J. R. (2005). Varieties of attention-deficit/hyperactivity disorder-related intra-individual variability. *Biol. Psychiatry* 57, 1416–1423. doi: 10.1016/j.biopsych.2004.12.005
- Chabris, C. F., Jerde, T. E., Woolley, A. W., Gerbasi, M. E., Schuldt, J. P., Bennett, S. L., et al. (2006). *Spatial and Object Visualization Cognitive Styles: Validation Studies In 3800 Individuals*. Technical Report No. 2. Cambridge, MA: Harvard University.
- Cipora, K., Hohol, M., Nuerk, H.-C., Willmes, K., Brożek, B., Kucharzyk, B., et al. (2016). Professional mathematicians differ from controls in their spatial-numerical associations. *Psychol. Res.* 80, 710–726. doi: 10.1007/s00426-015-0677-6
- Cipora, K., and Nuerk, H.-C. (2013). Is the SNARC effect related to the level of mathematics? No systematic relationship observed despite more power, more repetitions, and more direct assessment of arithmetic skill. *Q. J. Exp. Psychol.* 66, 1974–1991. doi: 10.1080/17470218.2013.772215
- Cipora, K., van Dijk, J.-P., Georges, C., Masson, N., Schiltz, C., Pesenti, M., et al. (2018). “Minority pulls the sample mean: only about 35% of individuals reveal a consistent SNARC effect,” in *Proceedings of the European Workshop on Cognitive Neuropsychology*, Brixen.

- Cipora, K., and Wood, G. (2012). Optimal power to detect (between group differences) in SNARC–Monte Carlo study. *Poster Presented on XXXth European Workshop on Cognitive Neuropsychology*, Brixen.
- Cipora, K., and Wood, G. (2017). Finding the SNARC instead of hunting it: a 20*20 Monte Carlo investigation. *Front. Psychol.* 8:1194. doi: 10.3389/fpsyg.2017.01194
- Cohen, J. D., Dunbar, K., and McClelland, J. L. (1990). On the control of automatic processes: a parallel distributed processing account of the Stroop effect. *Psychol. Rev.* 97, 332–361. doi: 10.1037/0033-295X.97.3.332
- Coles, M. G. (1989). Modern mind-brain reading: psychophysiology, physiology, and cognition. *Psychophysiology* 26, 251–269. doi: 10.1111/j.1469-8986.1989.tb01916.x
- Cook, R. D. (1979). Influential observations in linear regression. *J. Am. Stat. Assoc.* 74, 169–174. doi: 10.1080/01621459.1979.10481634
- Davelaar, E. J., and Stevens, J. (2009). Sequential dependencies in the Eriksen flanker task: a direct comparison of two competing accounts. *Psychon. Bull. Rev.* 16, 121–126. doi: 10.3758/PBR.16.1.121
- De Houwer, J. (2003). On the role of stimulus-response and stimulus-stimulus compatibility in the Stroop effect. *Mem. Cogn.* 31, 353–359. doi: 10.3758/BF03194393
- Dehaene, S., Bossini, S., and Giraux, P. (1993). The mental representation of parity and number magnitude. *J. Exp. Psychol. Gen.* 122, 371–396. doi: 10.1037/0096-3445.122.3.371
- Dehaene, S., and Brannon, E. (eds) (2011). *Space, Time and Number in the Brain*. Cambridge, MA: Academic Press.
- Dehaene, S., Dupoux, E., and Mehler, J. (1990). Is numerical comparison digital? Analogical and symbolic effects in two-number number comparison. *J. Exp. Psychol. Hum. Percept. Perform.* 16, 626–641. doi: 10.1037/0096-1523.16.3.626
- Eriksen, B. A., and Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target in a nonsearch task. *Percept. Psychophys.* 16, 143–149. doi: 10.3758/BF03203267
- Eriksen, C. W., and Schultz, D. W. (1979). Information processing in visual search: a continuous flow conception and experimental results. *Percept. Psychophys.* 25, 249–263. doi: 10.3758/BF03198804
- Falkenstein, M., Hielscher, H., Dziobek, I., Schwarzenau, P., Hoormann, J., Sunderman, B., et al. (2001). Action monitoring, error detection, and the basal ganglia: an ERP study. *Neuroreport* 12, 157–161. doi: 10.1097/00001756-200101220-00039
- Faul, F., Erdfelder, E., Buchner, A., and Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav. Res. Methods* 41, 1149–1160. doi: 10.3758/BRM.41.4.1149
- Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/BF03193146
- Fias, W., Brysbaert, M., Geypens, F., and d'Ydewalle, G. (1996). The importance of magnitude information in numerical processing: evidence from the SNARC effect. *Math. Cogn.* 2, 95–110. doi: 10.1080/135467996387552
- Fias, W., and van Dijck, J.-P. (2016). The temporary nature of number—space interactions. *Can. J. Exp. Psychol.* 70, 33–40. doi: 10.1037/cep0000071
- Fias, W., van Dijck, J.-P., and Gevers, W. (2011). “How is number associated with space? The role of working memory,” in *Space, Time and Number in the Brain*, eds S. Dehaene and E. Brannon (Cambridge, MA: Academic Press), 133–148.
- Fischer, M. H., and Shaki, S. (2014). Spatial associations in numerical cognition—From single digits to arithmetic. *Q. J. Exp. Psychol.* 67, 1461–1483. doi: 10.1080/17470218.2014.927515
- Georges, C. (2017). *Number-Space Associations as Indexed by the SNARC Effect – Their Relations to Mathematical Abilities and Anxiety & Their Underlying Cognitive Mechanisms*. Doctoral thesis, University of Luxembourg, Luxembourg.
- Georges, C., Hoffmann, D., and Schiltz, C. (2016). How math anxiety relates to number–space associations. *Front. Psychol.* 7:1401. doi: 10.3389/fpsyg.2016.01401
- Georges, C., Hoffmann, D., and Schiltz, C. (2017). How and why do number-space associations co-vary in implicit and explicit magnitude processing tasks? *J. Numer. Cogn.* 3, 182–211. doi: 10.5964/jnc.v3i2.46
- Gevers, W., Ratinckx, E., De Baene, W., and Fias, W. (2006). Further evidence that the SNARC effect is processed along a dual-route architecture: evidence from the lateralized readiness potential. *Exp. Psychol.* 53, 58–68. doi: 10.1027/1618-3169.53.1.58
- Gevers, W., Santens, S., Dhooge, E., Chen, Q., Van den Bossche, L., Fias, W., et al. (2010). Verbal-spatial and visuospatial coding of number-space interactions. *J. Exp. Psychol. Gen.* 139, 180–190. doi: 10.1037/a0017688
- Ginsburg, V., and Gevers, W. (2015). Spatial coding of ordinal information in short- and long-term memory. *Front. Hum. Neurosci.* 9:8. doi: 10.3389/fnhum.2015.00008
- Glisky, E. L. (2007). “Changes in cognitive function in human aging,” in *Brain Aging: Models, Methods, and Mechanisms*, ed. D. R. Riddle (Boca Raton, FL: CRC Press).
- Gloria, A. M., Castellanos, J., and Herrera, N. (2016). The reliability and validity of the cultural congruity and university environment scales with Chicana/o community college students. *Community Coll. J. Res. Pract.* 40, 426–438. doi: 10.1080/10668926.2015.1066276
- Goldfarb, L., and Henik, A. (2007). Evidence for task conflict in the Stroop effect. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 1170–1176. doi: 10.1037/0096-1523.33.5.1170
- Goodwin, L. D., and Leech, N. L. (2006). Understanding correlation: factors that affect the size of r. *J. Exp. Educ.* 74, 249–266. doi: 10.3200/JEXE.74.3.249-266
- Gratton, G., Coles, M. G., Sirevaag, E. J., Eriksen, C. W., and Donchin, E. (1988). Pre- and post-stimulus activation of response channels: a psychophysiological analysis. *J. Exp. Psychol. Hum. Percept. Perform.* 14, 331–344. doi: 10.1037/0096-1523.14.3.331
- Greene, D. J., Barnea, A., Herzberg, K., Rassis, A., Neta, M., Raz, A., et al. (2008). Measuring attention in the hemispheres: the lateralized attention network test (LANT). *Brain Cogn.* 66, 21–31. doi: 10.1016/j.bandc.2007.05.003
- Heitz, R. P., and Engle, R. W. (2007). Focusing the spotlight: individual differences in visual attention control. *J. Exp. Psychol. Gen.* 136, 217–240. doi: 10.1037/0096-3445.136.2.217
- Herrera, A., Macizo, P., and Semenza, C. (2008). The role of working memory in the association between number magnitude and space. *Acta Psychol.* 128, 225–237. doi: 10.1016/j.actpsy.2008.01.002
- Hoffmann, D., Mussolin, C., Martin, R., and Schiltz, C. (2014a). The impact of mathematical proficiency on the number-space association. *PLoS One* 9:e85048. doi: 10.1371/journal.pone.0085048
- Hoffmann, D., Pigat, D., and Schiltz, C. (2014b). The impact of inhibition capacities and age on number–space associations. *Cogn. Process.* 15, 329–342. doi: 10.1007/s10339-014-0601-9
- Huber, S., Klein, E., Moeller, K., and Willmes, K. (2016). Spatial-numerical and ordinal positional associations coexist in parallel. *Front. Psychol.* 7:438. doi: 10.3389/fpsyg.2016.00438
- Hutchison, K. A. (2011). The interactive effects of listwide control, item-based control, and working memory capacity on Stroop performance. *J. Exp. Psychol. Learn. Mem. Cogn.* 37, 851–860. doi: 10.1037/a0023437
- Ilán, A. B., and Polich, J. (1999). P300 and response time from a manual Stroop task. *Clin. Neurophysiol.* 110, 367–373. doi: 10.1016/S0168-5597(98)00053-7
- Jonides, J., and Nee, D. E. (2006). Brain mechanisms of proactive interference in working memory. *Neuroscience* 139, 181–193. doi: 10.1016/j.neuroscience.2005.06.042
- Jonides, J., Smith, E. E., Marshuetz, C., Koeppe, R. A., and Reuter-Lorenz, P. A. (1998). Inhibition in verbal working memory revealed by brain activation. *Proc. Natl. Acad. Sci. U.S.A.* 95, 8410–8413. doi: 10.1073/pnas.95.14.8410
- Kane, M. J., and Engle, R. W. (2003). Working memory capacity and the control of attention: the contributions of goal neglect, response competition, and task set to Stroop interference. *J. Exp. Psychol. Gen.* 132, 47–70. doi: 10.1037/0096-3445.132.1.47
- Kessler, R. C., Adler, L., Ames, M., Demler, O., Faraone, S., Hiripi, E., et al. (2005). The World Health Organization adult ADHD self-report scale (ASRS): a short screening scale for use in the general population. *Psychol. Med.* 35, 245–256. doi: 10.1017/S0033291704002892
- Keus, I. M., Jenks, K. M., and Schwarz, W. (2005). Psychophysiological evidence that the SNARC effect has its functional locus in a response selection stage. *Cogn. Brain Res.* 24, 48–56. doi: 10.1016/j.cogbrainres.2004.12.005
- Khng, K. H., and Lee, K. (2014). The relationship between Stroop and stop-signal measures of inhibition in adolescents: influences from variations in context

- and measure estimation. *PLoS One* 9:e101356. doi: 10.1371/journal.pone.0101356
- Kim, C., Kroger, J. K., and Kim, J. (2011). A functional dissociation of conflict processing within anterior cingulate cortex. *Hum. Brain Mapp.* 32, 304–312. doi: 10.1002/hbm.21020
- King, J. A., Colla, M., Bras, M., Heuser, I., and Von Cramon, D. Y. (2007). Inefficient cognitive control in adult ADHD: evidence from trial-by-trial Stroop test and cued task switching performance. *Behav. Brain Funct.* 3:42. doi: 10.1186/1744-9081-3-42
- Klein, G. S. (1964). Semantic power measured through the interference of words with color-naming. *Am. J. Psychol.* 77, 576–588. doi: 10.2307/1420768
- Kornblum, S., Hasbroucq, T., and Osman, A. (1990). Dimensional overlap: cognitive basis for stimulus-response compatibility—a model and taxonomy. *Psychol. Rev.* 97, 253–270. doi: 10.1037/0033-295X.97.2.253
- Kornblum, S., and Lee, J.-W. (1995). Stimulus-response compatibility with relevant and irrelevant stimulus dimensions that do and do not overlap with the response. *J. Exp. Psychol. Hum. Percept. Perform.* 21, 855–875. doi: 10.1037/0096-1523.21.4.855
- Kozhevnikov, M., Kosslyn, S., and Shephard, J. (2005). Spatial versus object visualizers: a new characterization of visual cognitive style. *Mem. Cogn.* 33, 710–726. doi: 10.3758/BF03195337
- La Heij, W. (1988). Components of Stroop-like interference in picture naming. *Mem. Cogn.* 16, 400–410. doi: 10.3758/BF03214220
- Lamm, C., Bauer, H., Vitouch, O., and Gstättner, R. (1999). Differences in the ability to process a visuo-spatial task are reflected in event-related slow cortical potentials of human subjects. *Neurosci. Lett.* 269, 137–140. doi: 10.1016/S0304-3940(99)00441-3
- Lansbergen, M. M., Kenemans, J. L., and van Engeland, H. (2007). Stroop interference and attention-deficit/hyperactivity disorder: a review and meta-analysis. *Neuropsychology* 21, 251–262. doi: 10.1037/0894-4105.21.2.251
- Laski, E. V., and Siegler, R. S. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child Dev.* 78, 1723–1743. doi: 10.1111/j.1467-8624.2007.01087.x
- Leung, H. C., Skudlarski, P., Gatenby, J. C., Peterson, B. S., and Gore, J. C. (2000). An event-related functional MRI study of the Stroop color word interference task. *Cereb. Cortex* 10, 552–560. doi: 10.1093/cercor/10.6.552
- Leuthold, H., Sommer, W., and Ulrich, R. (2004). Preparing for action: inferences from CNV and LRP. *J. Psychophysiol.* 18, 77–88. doi: 10.1027/0269-8803.18.23.77
- Link, T., Nuerk, H.-C., and Moeller, K. (2014). On the relation between the mental number line and arithmetic competencies. *Q. J. Exp. Psychol.* 67, 1597–1613. doi: 10.1080/17470218.2014.892517
- Long, D. L., and Prat, C. S. (2002). Working memory and Stroop interference: an individual differences investigation. *Mem. Cogn.* 30, 294–301. doi: 10.3758/BF03195290
- Lovejoy, D. W., Ball, J. D., Keats, M., Stutts, M. L., Spain, E., Janda, L., et al. (1999). Neuropsychological performance of adults with attention deficit hyperactivity disorder (ADHD): diagnostic classification estimates for measures of frontal lobe/executive functioning. *J. Int. Neuropsychol. Soc.* 5, 222–233. doi: 10.1017/S1355617799533055
- Lundervold, A. J., Adolphsdtottir, S., Halleland, H., Halmøy, A., Plessen, K., and Haavik, J. (2011). Attention network test in adults with ADHD — the impact of affective fluctuations. *Behav. Brain Funct.* 7, 27–34. doi: 10.1186/1744-9081-7-27
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: an integrative review. *Psychol. Bull.* 109, 163–203. doi: 10.1037/0033-2909.109.2.163
- MacLeod, J., Lawrence, M., McConnell, M., Eskes, G., Klein, R., and Shore, D. (2010). Appraising the ANT: psychometric and theoretical considerations of the attention network test. *Neuropsychology* 24, 637–651. doi: 10.1037/a0019803
- Matza, L. S., Van Brunt, D. L., Cates, C., and Murray, L. T. (2011). Test-retest reliability of two patient-report measures for use in adults with ADHD. *J. Atten. Disord.* 15, 557–563. doi: 10.1177/1087054710372488
- Melcher, T., and Gruber, O. (2009). Decomposing interference during Stroop performance into different conflict factors: an event-related fMRI study. *Cortex* 45, 189–200. doi: 10.1016/j.cortex.2007.06.004
- Milham, M. P., Banich, M. T., Webb, A., Barad, V., Cohen, N. J., Wszalek, T., et al. (2001). The relative involvement of anterior cingulate and prefrontal cortex in attentional control depends on nature of conflict. *Brain Res. Cogn. Brain Res.* 12, 467–473. doi: 10.1016/S0926-6410(01)00076-3
- Moyer, R. S., and Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature* 215, 1519–1520. doi: 10.1038/2151519a0
- Muchinsky, P. M. (1996). The correction for attenuation. *Educ. Psychol. Meas.* 56, 63–75. doi: 10.1177/0013164496056001004
- Nee, D. E., Wager, T. D., and Jonides, J. (2007). Interference resolution: insights from a meta-analysis of neuroimaging tasks. *Cogn. Affect. Behav. Neurosci.* 7, 1–17. doi: 10.3758/CABN.7.1.1
- Nieuwenhuis, S., Ridderinkhof, K. R., Talsma, D., Coles, M. G., Holroyd, C. B., Kok, A., et al. (2002). A computational account of altered error processing in older age: dopamine and the error-related negativity. *Cogn. Affect. Behav. Neurosci.* 2, 19–36. doi: 10.3758/CABN.2.1.19
- Nigg, J. T., Willcutt, E. G., Doyle, A. E., and Sonuga-Barke, E. J. S. (2005). Causal heterogeneity in attention-deficit/hyperactivity disorder: Do we need neuropsychologically impaired subtypes? *Biol. Psychiatry* 57, 1224–1230. doi: 10.1016/j.biopsych.2004.08.025
- Ninaus, M., Moeller, K., Kaufmann, L., Fischer, M. H., Nuerk, H.-C., and Wood, G. (2017). Cognitive mechanisms underlying directional and non-directional spatial-numerical associations across the lifespan. *Front. Psychol.* 8:1421. doi: 10.3389/fpsyg.2017.01421
- Nuerk, H., Bauer, F., Krummenacher, J., Heller, D., and Willmes, K. (2005). The power of the mental number line: how the magnitude of unattended numbers affects performance in an Eriksen task. *Psychol. Sci.* 47, 34–50.
- Osborne, J. W., and Waters, E. (2002). Four assumptions of multiple regression that researchers should always test. *Pract. Assess. Res. Eval.* 8, 1–5.
- Pearson, K., and Filon, L. N. G. (1898). Mathematical contributions to theory of evolution: IV. On the probable errors of frequency constants and on the influence of random selection and correlation. *Philos. Trans. R. Soc. Lond. A* 191, 229–311. doi: 10.1098/rsta.1898.0007
- Priftis, K., Zorzi, M., Meneghello, F., Marenzi, R., and Umiltà, C. (2006). Explicit versus implicit processing of representational space in neglect: dissociations in accessing the mental number line. *J. Cogn. Neurosci.* 18, 680–688. doi: 10.1162/jocn.2006.18.4.680
- Proctor, R. W., and Cho, Y. S. (2006). Polarity correspondence: a general principle for performance of speeded binary classification tasks. *Psychol. Bull.* 132, 416–442. doi: 10.1037/0033-2909.132.3.416
- Rapport, L. J., Van Voorhis, A., Tzelepis, A., and Friedman, S. R. (2001). Executive functioning in adult attention-deficit hyperactivity disorder. *Clin. Neuropsychol.* 15, 479–491. doi: 10.1076/clin.15.4.479.1878
- Redick, T. S., and Engle, R. W. (2006). Working memory capacity and Attention Network Test performance. *Appl. Cogn. Psychol.* 20, 713–721. doi: 10.1002/acp.1224
- Restle, F. (1970). Speed of adding and comparing numbers. *J. Exp. Psychol.* 83(2 Pt 1), 274–278. doi: 10.1037/h0028573
- Sahinoglu, B., and Dogan, G. (2016). Event-related potentials and the Stroop effect. *Eurasian J. Med.* 48, 53–57. doi: 10.5152/eurasianjmed.2016.16012
- Schmidt, J. R., and Cheesman, J. (2005). Dissociating stimulus-stimulus and response-response effects in the Stroop task. *Can. J. Exp. Psychol.* 59, 132–138. doi: 10.1037/h0087468
- Seidman, L. J. (2006). Neuropsychological functioning in people with ADHD across the lifespan. *Clin. Psychol. Rev.* 26, 466–485. doi: 10.1016/j.cpr.2006.01.004
- Sergeant, J. A., Geurts, H., and Oosterlaan, J. (2002). How specific is a deficit of executive functioning for attention-deficit/hyperactivity disorder? *Behav. Brain Res.* 130, 3–28.
- Sharma, D., and McKenna, F. P. (1998). Differential components of the manual and vocal Stroop tasks. *Mem. Cogn.* 26, 1033–1040. doi: 10.3758/BF03201181
- Spearman, C. (1904). The proof and measurement of association between two things. *Am. J. Psychol.* 15, 72–101. doi: 10.2307/1412159
- Spearman, C. (1910). Correlation calculated from faulty data. *Br. J. Psychol.* 3, 271–295. doi: 10.1111/j.2044-8295.1910.tb00206.x
- Stins, J. F., Polderman, J. C., Boomsma, D. I., and de Geus, E. J. (2005). Response interference and working memory in 12-year-old children. *Child Neuropsychol.* 11, 191–201. doi: 10.1080/092970409011351

- Stins, J. F., van Baal, G. C. M., Polderman, T. J., Verhulst, F. C., and Boomsma, D. I. (2004). Heritability of Stroop and flanker performance in 12-year old children. *BMC Neurosci.* 5:49. doi: 10.1186/1471-2202-5-49
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *J. Exp. Psychol.* 18, 643–662. doi: 10.1037/h0054651
- Szucs, D., and Soltész, F. (2010). Stimulus and response conflict in the color–word Stroop task: a combined electro-myography and event-related potential study. *Brain Res.* 1325, 63–76. doi: 10.1016/j.brainres.2010.02.011
- Szucs, D., Soltész, F., Bryce, D., and Whitebread, D. (2009). Real-time tracking of motor response activation and response competition in a Stroop task in young children: a lateralized readiness potential study. *J. Cogn. Neurosci.* 21, 2195–2206. doi: 10.1162/jocn.2009.21220
- Valle-Inclán, F. (1996). The locus of interference in the Simon effect: an ERP study. *Biol. Psychol.* 43, 147–162. doi: 10.1016/0301-0511(95)05181-3
- Van der Elst, W., Van Boxtel, M. P. J., Van Breukelen, G. J. P., and Jolles, J. (2006). The Stroop color-word test: influence of age, sex, and education; and normative data for a large sample across the adult age range. *Assessment* 13, 62–79. doi: 10.1177/1073191105283427
- van Dijck, J.-P., and Fias, W. (2011). A working memory account for spatial-numerical associations. *Cognition* 119, 114–119. doi: 10.1016/j.cognition.2010.12.013
- van Dijck, J.-P., Gevers, W., and Fias, W. (2009). Numbers are associated with different types of spatial information depending on the task. *Cognition* 113, 248–253. doi: 10.1016/j.cognition.2009.08.005
- van Galen, M. S., and Reitsma, P. (2008). Developing access to number magnitude: a study of the SNARC effect in 7- to 9-year-olds. *J. Exp. Child Psychol.* 101, 99–113. doi: 10.1016/j.jecp.2008.05.001
- van Veen, V., and Carter, C. S. (2005). Separating semantic conflict and response conflict in the Stroop task: a functional MRI study. *Neuroimage* 27, 497–504. doi: 10.1016/j.neuroimage.2005.04.042
- Vaurio, R. G., Simmonds, D. J., and Mostofsky, S. H. (2009). Increased intra-individual reaction time variability in attention-deficit/hyperactivity disorder across response inhibition tasks with different cognitive demands. *Neuropsychologia* 47, 2389–2396. doi: 10.1016/j.neuropsychologia.2009.01.022
- Viarouge, A., Hubbard, E. M., and McCandliss, B. D. (2014). The cognitive mechanisms of the SNARC effect: an individual differences approach. *PLoS One* 9:e95756. doi: 10.1371/journal.pone.0095756
- Walker, A. J., Shores, E. A., Trollor, J. N., Lee, T., and Sachdev, P. S. (2000). Neuropsychological functioning of adults with attention deficit hyperactivity disorder. *J. Clin. Exp. Neuropsychol.* 22, 115–124. doi: 10.1076/1380-3395(200002)22:1;1-8;FT115
- Wendt, M., Luna-Rodriguez, A., and Jacobsen, T. (2012). Conflict-induced perceptual filtering. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 675–686. doi: 10.1037/a0025902
- West, R., and Alain, C. (2000). Age-related decline in inhibitory control contributes to the increased Stroop effect observed in older adults. *Psychophysiology* 37, 179–189. doi: 10.1111/1469-8986.3720179
- Wood, G., Willmes, K., Nuerk, H.-C., and Fischer, M. (2008). On the cognitive link between space and number: a meta-analysis of the SNARC effect. *Psychol. Sci. Q.* 50, 489–525.
- Zhang, H., and Kornblum, S. (1998). The effects of stimulus-response mapping and irrelevant stimulus-response and stimulus-stimulus overlap in four-choice Stroop tasks with single-carrier stimuli. *J. Exp. Psychol. Hum. Percept. Perform.* 24, 3–19. doi: 10.1037/0096-1523.24.1.3
- Zhang, H. H., Zhang, J., and Kornblum, S. (1999). A parallel distributed processing model of stimulus-stimulus and stimulus-response compatibility. *Cogn. Psychol.* 38, 386–432. doi: 10.1006/cogp.1998.0703
- Zhu, D. C., Zacks, R. T., and Slade, J. M. (2010). Brain activation during interference resolution in young and older adults: an fMRI study. *Neuroimage* 50, 810–817. doi: 10.1016/j.neuroimage.2009.12.087
- Zmigrod, S., Zmigrod, L., and Hommel, B. (2016). Transcranial direct current stimulation (tDCS) over the right dorsolateral prefrontal cortex affects stimulus conflict but not response conflict. *Neuroscience* 322, 320–325. doi: 10.1016/j.neuroscience.2016.02.046
- Zorzi, M., Bonato, M., Treccani, B., Scalabrini, G., Marenzi, R., and Priftis, K. (2012). Neglect impairs explicit processing of the mental number line. *Front. Hum. Neurosci.* 6:125. doi: 10.3389/fnhum.2012.00125
- Zurrón, M., Pouso, M., Lindín, M., Galdó, S., and Díaz, F. (2009). Event-related potentials with the Stroop colour-word task: timing of semantic conflict. *Int. J. Psychophysiol.* 72, 246–252. doi: 10.1016/j.ijpsycho.2009.01.002

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Georges, Hoffmann and Schiltz. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Connection Between Spatial and Mathematical Ability Across Development

Christopher J. Young^{1*}, Susan C. Levine¹ and Kelly S. Mix²

¹ Department of Psychology, University of Chicago, Chicago, IL, United States, ² Department of Human Development and Quantitative Methodology, University of Maryland, College Park, MD, United States

In this article, we review approaches to modeling a connection between spatial and mathematical thinking across development. We critically evaluate the strengths and weaknesses of factor analyses, meta-analyses, and experimental literatures. We examine those studies that set out to describe the nature and number of spatial and mathematical skills and specific connections between these abilities, especially those that included children as participants. We also find evidence of strong spatial-mathematical connections and transfer from spatial interventions to mathematical understanding. Finally, we map out the kinds of studies that could enhance our understanding of the mechanisms by which spatial and mathematical processing are connected and the principles by which mathematical outcomes could be enhanced through spatial training in educational settings.

OPEN ACCESS

Edited by:

Hans-Christoph Nuerk,
Universität Tübingen, Germany

Reviewed by:

Robert S. Siegler,
Carnegie Mellon University,
United States

Zachary Hawes,
University of Western Ontario, Canada
Victoria Simms,
Ulster University, United Kingdom

*Correspondence:

Christopher J. Young
youngcj@uchicago.edu

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 01 February 2018

Accepted: 30 April 2018

Published: 04 June 2018

Citation:

Young CJ, Levine SC and Mix KS
(2018) The Connection Between
Spatial and Mathematical Ability
Across Development.
Front. Psychol. 9:755.
doi: 10.3389/fpsyg.2018.00755

Keywords: spatial cognition, mathematical concepts, factor analysis, statistical, developmental psychology, process modeling

INTRODUCTION

Spatial ability contributes to performance in science, technology, engineering, and mathematics (STEM) domains even controlling for verbal and mathematical abilities (Shea et al., 2001; Wai et al., 2009). In addition, spatial reasoning task performance has been found to correlate with mathematical task performance (e.g., Dehaene et al., 1999), suggesting that spatial reasoning skills overlap with, and could be necessary for, mathematical reasoning skills (Tosto et al., 2014). One correlation supported by cognitive and developmental research is between representations of numerical and spatial magnitudes. Spatial skills have been found to correlate with numerical magnitude representations across broad age ranges, from preschoolers (Gunderson et al., 2012) to adults (Sella et al., 2016). Further, spatial and numerical magnitude representations have overlapping neural representations (Piazza et al., 2007; Holloway et al., 2010). In this article, we review evidence for the connections between spatial and mathematical skills across development that has been gleaned from factor analyses, meta-analyses, and experimentation. We then suggest productive ways to elucidate spatial-mathematical connections and discuss ways that modeling could be used to improve mathematics learning.

FACTOR ANALYSIS

Both spatial and mathematical ability have been investigated since the early days of psychological science using factor analytical methods that sought to map the “structure of the intellect” (Spearman, 1927; Thurstone, 1938). This research showed a connection between spatial and mathematical domains, yet the mechanisms by which training

spatial thinking can promote mathematical thinking are still not well understood. Across various factor analyses of spatial skills that have been conducted in adults, the most consistent finding is that there are multiple spatial skills, such as spatial visualization (imagining transformations) and spatial relations and spatial orientation (perceiving object position and angle) (Michael et al., 1957; McGee, 1979; Lohman, 1988; Carroll, 1993). Factor analyses carried out on mathematical measures over various ages have revealed latent factors that do not appear to be specific to mathematics (e.g., deductive reasoning and adaptability to a new task among 10th grade students, Kline, 1960; abstraction, analysis, application among elementary school students Rusch, 1957). These studies are notable in that some theorists have found evidence of a spatial factor in mathematics (e.g., Kline, 1960; Werdelin, 1966) and others have argued that there is a spatial sensorimotor intelligence factor important to mathematical reasoning (Coleman, 1960; Skemp, 1961; Aiken, 1970).

Separate but Correlated Spatial and Mathematical Thinking Factors

While many studies have found evidence of connections between spatial and numerical tasks in young children, only recently have studies explored the factor structure of their spatial and mathematical skills. Mix et al. (2016, 2017a) have used factor analyses to examine the connections among a broad range of mathematical and spatial tasks in elementary school age children. Mix et al. (2016) administered a battery of tasks that had the greatest likelihood of showing spatial-mathematical connections based on the literature, including connections between (1) spatial visualization and complex mathematical relations, (2) form perception and symbolic reasoning, and (3) spatial scaling and numerical estimation (Landy and Goldstone, 2010; Slusser et al., 2013; Thompson et al., 2013, respectively). These tasks were included in order to identify which underlying variables that connect spatial and mathematical domains in kindergarten, third and sixth grades.

Between kindergarten and sixth grade range, all spatial tasks loaded together on a distinctly spatial factor, and all mathematical tasks loaded on a distinctly mathematical factor (Mix et al., 2016, 2017a). However, there was a moderate correlation between the two factors ($r_s = 0.50\text{--}0.53$), even when controlling for verbal ability, suggesting that although the spatial and mathematical domains are distinct, there is a significant relation between these domains. Even though verbal ability accounted for a significant portion of variance in mathematical skills in each grade tested, spatial skills accounted for a greater proportion of variance (Mix et al., 2016). Cross-loadings between the spatial and mathematical factors and tasks in the two domains also indicate specific connections. In kindergarteners, mental rotation was significantly related to the mathematical factor, whereas in sixth graders visuospatial working memory and form copying were significantly related to the mathematical factor. One possible explanation for the change in cross-loadings over development is that mathematical thinking relies at first on dynamic, object-focused spatial processes (mental rotation) and later on more

static, memory-related spatial processes (visuospatial working memory and visuomotor integration).

Strengths and Limitations of Factor Analysis Evidence

Factor analysis is a useful tool for isolating the source of correlations and removing measurement error (Bollen, 1989) as well as for testing competing theories (Gerbing and Hamilton, 1996; Tomarken and Waller, 2005). However, factor analysis requires a large number of participants over a breadth of tasks in a domain to achieve a stable structure (Hair et al., 1995; MacCallum et al., 1999). The biggest limitation of factor analysis lies in the theorist; interpretation of results is a large part of proper factor analysis because the results do not uniquely point to any single interpretation of the meaning of the underlying latent variables that are revealed (Armstrong and Soelberg, 1968; Rummel, 1970). Thus, when relations do emerge from factor analysis, other methods must be used to establish mechanisms underlying these relations.

META-ANALYTIC AND EXPERIMENTAL STUDIES

In addition to factor analyses, researchers have tackled the question of how the domains of space and math are connected through targeted experimental studies and meta-analyses. In this section, we outline prominent theories about the divisions in each domain and evidence for correlations between spatial and mathematical skills. Understanding these theories is important because they can help us to understand which particular facet or type of spatial thinking is linked to a particular type of mathematical thinking.

One comprehensive meta-analysis of spatial skills training by Uttal et al. (2013) assumed a 2×2 typology supported by behavioral (Newcombe and Shipley, 2015) and neurological evidence (e.g., Chatterjee, 2008). Specifically, relations between objects are processed differently than relations of feature within an object (the extrinsic-intrinsic division). Further, spatial information conveyed by a static viewing of objects and scenes is processed differently than movements and transformations of these objects and scenes (the static-dynamic division). In their factor analysis testing the 2×2 typology, (Mix et al., under review) found evidence for distinct spatial factors for tasks involving within object (intrinsic) vs. between object (extrinsic) information, but did not find support for spatial tasks separating according to the static-dynamic distinction (Mix et al. under review). Echoing this finding, Kozhenikov et al. found evidence that some children process spatial information intrinsic to objects better (object visualizers) whereas others process spatial information that involves between object relationships better (spatial visualizers) but did not find that these groups of children differed in their ability to process dynamic and static imagery (Kozhevnikov et al., 2005).

The number and nature of basic mathematical skills that underlie mathematical thinking are also in question. For example, a distinction has been made between core number

systems that represent exact and approximate number (Carey, 2004; Feigenson et al., 2004), between core systems for approximate number and ratio (Matthews and Hubbard, 2017), and between core approximate number system and exact number ability enabled by symbolic knowledge (e.g., Carey, 2004). However, the debate about the systems that characterize mathematical thinking has taken on a more pragmatic turn than those concerning spatial thinking. For instance, there are direct educational implications to whether core mathematical skills facilitate later symbolic mathematical understanding and achievement and how the latter might affect the former (e.g., Feigenson et al., 2013; Schneider et al., 2017) or whether mathematics is better taught through concepts or procedures (e.g., Schoenfeld, 1985), or abstractly or concretely (e.g., Kaminski et al., 2009). Researchers also debate which kinds of early mathematical skills relate to later mathematical achievement (e.g., understanding patterns, Rittle-Johnson et al., 2017; thinking symbolically, Schneider et al., 2017, or one's ordinal vs. absolute sense of number Lyons et al., 2014). These debates raise interesting questions about the connection between spatial skills, early mathematical skills, and later mathematical achievement. For example, does a particular type of spatial skill relate to children's ability to learn particular early mathematical skills more quickly, and are these the early mathematical skills that relate most strongly to later mathematical achievement?

What Skills Are Used in Both Spatial and Mathematical Problems?

Certain connections between specific spatial skills and mathematical skills have been observed (e.g., visuospatial working memory and computation, Raghubar et al., 2010) whereas others have not (e.g., between disembedding shapes from scenes and parsing information in charts, Clark, 1988) with little explanation as to why this is the case (for a review of these connections see Mix and Cheng, 2012). One frequently observed connection is between mental rotation and various math skills, across age and development and with a variety of different mental rotation task characteristics (Table 1). However, little is known about the processes that account for this connection, or whether there are other spatial-mathematical connections that may be even stronger. Thus, this correlational type of evidence fails to provide support for the theory that certain specific spatial skills are particularly important for mathematics achievement nor how they enable better performance and learning of specific mathematical skills. Answers to these questions are of high importance to successfully incorporating spatial learning into mathematical curricula.

Moving beyond correlational studies, studies that have measured the impact of training mental rotation on specific mathematical skills, have not yielded consistent findings, with some finding evidence of transfer (e.g., Cheng and Mix, 2014; Lowrie et al., 2017) and some not finding such evidence (Hawes et al., 2015b; Xu and LeFevre, 2016). There is little explanation, and as of yet no meta-analysis, to compare these cross-domain training studies or determine the overall effectiveness of training any individual spatial skills to improve mathematical

reasoning. In the next section, we argue that modeling and testing the processes involved in performing specific spatial and mathematical tasks can help us understand the connections between these two domains.

COGNITIVE PROCESS MODELS

Cognitive process models provide an account of the mental processes engaged when performing a specific task. What cognitive process or processes actually drive performance on a spatial task? Answering this question would also allow us to understand the mechanism that accounts for the connection between spatial skills, like mental rotation, and performance on mathematical tasks such as missing term problems (Cheng and Mix, 2014). This in turn would inform educational efforts to improve spatial thinking in ways that would be most helpful to mathematical thinking.

What is known about the processes used for spatial skills? Various studies have supported substantive divisions between particular kinds of spatial skills, e.g., the intrinsic-extrinsic divide separating tasks such as mental rotation from perspective taking (Huttenlocher and Presson, 1973; Kozhevnikov and Hegarty, 2001). However, studies with kindergarten through sixth grade children also show a great deal of overlap among a wide range of spatial skills (Mix et al., 2016, 2017a). Further, certain spatial skills, notably mental rotation and visuospatial working memory, have been found to cross-load onto a mathematical factor at particular grade levels. An important next step is to examine process models of spatial skills and how they are manifested (or not) on mathematical tasks, as illustrated below regarding mental rotation.

A Process View of Mental Rotation

Mental rotation was first described based on the finding that time to simulate the rotation of an object was related to the angle through which the object was rotated (Shepard and Metzler, 1971). Cognitive process models, supported by empirical studies, reveal that mental rotation actually involves multiple, non-obvious sub-components. Behavior is best fit by a model that involves carrying out small, successive, variable transformations, rather than a single rotation (Provost and Heathcote, 2015) and empirical work suggests that individuals actually rotate just one part of the object rather than all parts of the whole object (Xu and Franconeri, 2015). Further, modeling shows that the type of mental rotation problem influences the process that is engaged; when rotating complex stimuli, participants tend to be slower (Bethell-Fox and Shepard, 1988; Shepard and Metzler, 1988), which has been fit by computational models of mental rotation where task relevant features of the object are focused on and task irrelevant features are ignored (Lovett and Schultheis, 2014). Participants also frequently err in problems with complex stimuli by selecting the mirror image of the correct choice that is rotated to the same degree as the correct choice (e.g., among children Hawes et al., 2015a,b), a pattern of data that is explained by a model that parameterizes "confusability" between the target and its mirror (e.g., confusing a "d" for a "b," Kelley et al., 2000). Relatedly participants tend to use a fast flipping transformation

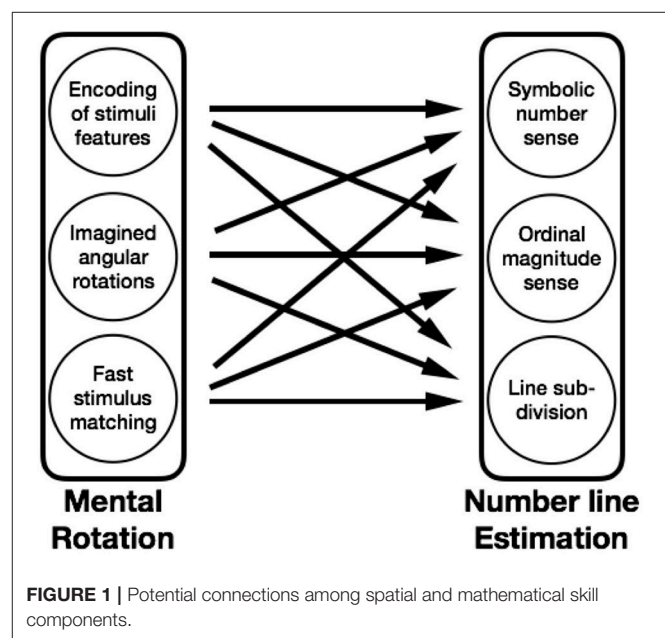
TABLE 1 | Observed relations between mental rotation and mathematical skills.

References	Average age	Mental rotation task characteristics	Measure or mathematical skill
Gunderson et al., 2012	5.4	Children's mental transformation task, 2D figure rotation/construction (Levine et al., 2018)	Number line estimation, approximate calculation
Kyttälä et al., 2003	6.16	Novell "Troll" task, 2D, same/different choice	General math skill
Carr et al., 2008	7.5	Cube rotation (Vandenberg and Kuse, 1978)	Arithmetic
Battista, 1990	12	Purdue spatial visualization Test, 3D images rotation (Guay, 1976)	Logical reasoning, geometric knowledge and problem solving
Hegarty and Kozhevnikov, 1999	12.08	Primary mental abilities, 2D rotation/figure completion (Thurstone, 1938)	Problem solving skill
Delgado and Prieto, 2004	13	Cube rotation (Peters, 1995)	Geometry, word problems
Casey et al., 1997	13.8	Cube rotation (Vandenberg and Kuse, 1978)	Geometry, SAT math
Kyttälä and Lehto, 2008	15.5	Cube rotation (Vandenberg and Kuse, 1978)	Mental arithmetic, geometry, word problems
Reuhkala, 2001	15.5	Cube rotation (Vandenberg and Kuse, 1978)	Math skill (mental arithmetic, algebra, geometry)
Geary et al., 2000	19	Cube rotation (Vandenberg and Kuse, 1978)	Arithmetic
Thompson et al., 2013	21.26	Cube rotation (Peters, 1995)	Compatibility effect of number comparison

akin to matching features for simple, 2D stimuli, which models of mental rotation have taken this into account (Kung and Hamm, 2010; Searle and Hamm, 2012). The varied components described by these models make clear that mental rotation is not a simple process, and that there are many steps needed to succeed at a mental rotation task.

Each of these modeled components of mental rotation performance has a potential role to play in the observed relationship between mental rotation and various mathematical skills over the course of development. If spatial constructs are actually based on wide-ranging processes it opens up the hypothesis space to determine the source of connections between spatial and mathematical thinking. Rather than a simple connection between two monolithic skills, there are numerous possible connections based on the components of each, and possibly even multiple ways a spatial skill can act in a single math problem. The work of figuring out which components are critical to the observed relation between spatial and mathematical skills, while daunting, is needed in order to unpack what otherwise are opaque connections.

To take one example, Gunderson et al. (2012) observed a predictive relationship between young children's mental transformation skill and their number line estimation. Individual differences in mental rotation performance could have arisen as a difference in any of the components identified above: the ability to carry out rotations, to focus on relevant spatial information, or to carry out non-rotational stimulus matching. Similarly, the number line estimation task, where participants are asked to determine the position of a number along a labeled line, could be decomposed into several components as well (e.g., accessing a representation of a number's magnitude when cued by its symbol, ordering those magnitudes precisely on a continuous number line, spatially subdividing the line at salient landmarks, Siegler and Opfer, 2003). Any or all of these components might be the source of the connection between number line estimations and spatial skill (see **Figure 1**). By designing studies that control for and model the components of both spatial and mathematical tasks, it should be possible to identify and



understand the mechanisms that explain links between spatial and mathematical thinking. This approach compliments and enriches the work focused on looking at the latent structure of skills, while not dwelling on an explanation of any one task but focusing on explaining important connections between latent skills.

EDUCATIONAL IMPLICATIONS

Meta-analyses provide strong evidence that training spatial skills in the laboratory result in significant improvements and transfer to other spatial skills (Uttal et al., 2013). However, evidence is more mixed about training spatial skills to improve mathematical skills (e.g., Cheng and Mix, 2014;

Hawes et al., 2015b; Simons et al., 2016; Lowrie et al., 2017). Broader training regimes in and out of the classroom have helped to improve mathematics performance in multiple age groups (e.g., Witt, 2011; Sorby et al., 2013; Bruce and Hawes, 2015), and more generally, spatial thinking has been shown to be a significant predictor of STEM outcomes, even controlling for mathematical and verbal thinking (Wai et al., 2009).

One finding substantiated by factor analyses and interventions is that spatial skills are more closely related to novel mathematical and scientific content than to STEM skills that are more familiar (Stieff, 2013; Mix et al., 2016), suggesting that it may be particularly important to provide students with spatial scaffolding when students are learning a new mathematical concept. Another set of findings suggests that providing students with a repertoire of spatial tools, such as gesture, rich spatial language, diagrams, and spatial analogies, (Newcombe, 2010; Levine et al., 2018) can facilitate their spatial thinking. Moreover, these tools, as well as 3-D manipulatives (Mix, 2010) have been found to facilitate learning mathematical concepts (e.g., Richland et al., 2012; Verdine et al., 2014; Hawes et al., 2017; Mix et al., 2017b). An overarching principle to guide the use of spatial thinking and tools in education is that supporting spatial thinking and learning beginning early in life may result in improvements in mathematics understanding, based on the general connection between spatial and mathematical factors as well as evidence that training particular spatial skills shows some transfer to mathematics skills. A promising avenue for future work is not just to support spatial thinking in general, but to show students how they can use this kind of thinking to solve particular kinds of mathematical problems (Casey, 2004).

REFERENCES

- Aiken, L. R. Jr. (1970). Attitudes toward mathematics. *Rev. Educ. Res.* 40, 551–596. doi: 10.3102/00346543040004551
- Armstrong, J. S., and Soelberg, P. (1968). On the interpretation of factor analysis. *Psychol. Bull.* 70:361. doi: 10.1037/h0026434
- Battista, M. T. (1990). Spatial visualization and gender differences in high school geometry. *J. Res. Math. Educ.* 47–60.
- Bethell-Fox, C. E., and Shepard, R. N. (1988). Mental rotation: effects of stimulus complexity and familiarity. *J. Exp. Psychol. Hum. Percept. Perform.* 14:12. doi: 10.1037/0096-1523.14.1.12
- Bollen, K. A. (1989). A new incremental fit index for general structural equation models. *Sociol. Methods Res.* 17, 303–316. doi: 10.1177/0049124189017003004
- Bruce, C. D., and Hawes, Z. (2015). The role of 2D and 3D mental rotation in mathematics for young children: what is it? Why does it matter? And what can we do about it? *ZDM* 47, 331–343. doi: 10.1007/s11858-014-0637-4
- Carey, S. (2004). Bootstrapping and the origin of concepts. *Daedalus* 133, 59–68. doi: 10.1162/001152604772746701
- Carr, M., Steiner, H. H., Kyser, B., and Biddlecomb, B. (2008). A comparison of predictors of early emerging gender differences in mathematics competency. *Learn. Individ. Diff.* 18, 61–75.
- Carroll, J. B. (1993). *Human Cognitive Abilities: A Survey of Factor-Analytic Studies*. Cambridge University Press.
- Casey, M. B., Nuttall, R. L., and Pizaris, E. (1997). Mediators of gender differences in mathematics college entrance test scores: a comparison of spatial skills with internalized beliefs and anxieties. *Dev. Psychol.* 33:669.

CONCLUSIONS

In this review, we critically evaluate the contributions of the factor analytic method to identifying and elucidating the connection between spatial and mathematical thinking across development. We highlighted a central gap in our knowledge—understanding the mechanisms connecting spatial and mathematical skills—which can be better addressed through targeted experimental studies that are informed by process models than by factor analytic studies. The findings that can emerge from this approach are important for increasing our basic understanding of why spatial and mathematical thinking are connected. They also hold promise for informing educational efforts to increase mathematical achievement by strengthening spatial thinking by training spatial skills, by encouraging the use of spatial tools, and by showing children how they can deploy these skills and tools to solve particular kind of mathematical problems.

AUTHOR CONTRIBUTIONS

CY wrote the original draft and led efforts to refine subsequent drafts for this article. All authors worked on a related chapter; SL and KM contributed substantially to the writing and editing of this article.

FUNDING

This research was supported by Institute of Education Sciences Grant R305A120416 to KM and SL, as well as National Science Foundation Spatial Intelligence Learning Center Grants SBE-1041707 and SBE-0541957 to SL.

- Casey, B. (2004). “Mathematics problem-solving adventures: a language-arts based supplementary series for early childhood that focuses on spatial sense,” in *Engaging Young Children In Mathematics: Results of the Conference on Standards for Pre-School and Kindergarten Mathematics Education*, eds D. Clements, J. Sarama, and M. A. DiBaise (Mahwah, NJ: Erlbaum Associates), 377–389.
- Chatterjee, A. (2008). “The neural organization of spatial thought and language,” in *Seminars in Speech and Language Vol. 29* (Thieme Medical Publishers), 226–238.
- Cheng, Y. L., and Mix, K. S. (2014). Spatial training improves children’s mathematics ability. *J. Cogn. Dev.* 15, 2–11. doi: 10.1080/15248372.2012.725186
- Clark, H. T. III. (1988). “Cognitive style and competence in mathematics problem solving,” in *Paper Presented at the Annual Meeting of the Eastern Educational Research Association*. Miami, FL.
- Coleman, R. H. (1960). *An analysis of Certain Components of Mathematical Ability and an attempt to Predict Mathematical Achievement in a Specific Situation*. Doctoral dissertation, Indiana University. University Microfilms, Ann Arbor.
- Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., and Tsivkin, S. (1999). Sources of mathematical thinking: behavioral and brain-imaging evidence. *Science* 284, 970–974. doi: 10.1126/science.284.5416.970
- Delgado, A. R., and Prieto, G. (2004). Cognitive mediators and sex-related differences in mathematics. *Intelligence* 32, 25–32.
- Feigenson, L., Dehaene, S., and Spelke, E. (2004). Core systems of number. *Trends Cogn. Sci.* 8, 307–314. doi: 10.1016/j.tics.2004.05.002

- Feigenson, L., Libertus, M. E., and Halberda, J. (2013). Links between the intuitive sense of number and formal mathematics ability. *Child Dev. Perspect.* 7, 74–79. doi: 10.1111/cdep.12019
- Geary, D. C., Saults, S. J., Liu, F., and Hoard, M. K. (2000). Sex differences in spatial cognition, computational fluency, and arithmetical reasoning. *J. Exp. Child Psychol.* 77, 337–353. doi: 10.1006/jecp.2000.2594
- Gerbing, D. W., and Hamilton, J. G. (1996). Viability of exploratory factor analysis as a precursor to confirmatory factor analysis. *Struct. Equ. Model. Mult. J.* 3, 62–72. doi: 10.1080/10705519609540030
- Guay, R. B. (1976). *Purdue Spatial Visualization Test*. West Lafayette, IN: Purdue Research Foundation.
- Gunderson, E. A., Ramirez, G., Beilock, S. L., and Levine, S. C. (2012). The relation between spatial skill and early number knowledge: the role of the linear number line. *Dev. Psychol.* 48, 1229–1241. doi: 10.1037/a0027433
- Hair, J. F. Jr., Anderson, R. E., Tatham, R. L., and William, C. (1995). *Multivariate Data Analysis with Readings*. Englewood Cliffs, NJ: Prentice Hall.
- Hawes, Z., LeFevre, J. A., Xu, C., and Bruce, C. D. (2015a). Mental rotation with tangible three-dimensional objects: a new measure sensitive to developmental differences in 4- to 8-year-old children. *Mind Brain Educ.* 9, 10–18. doi: 10.1111/mbe.12051
- Hawes, Z., Moss, J., Caswell, B., and Poliszczuk, D. (2015b). Effects of mental rotation training on children's spatial and mathematics performance: a randomized controlled study. *Trends Neurosci. Educ.* 4, 60–68. doi: 10.1016/j.tine.2015.05.001
- Hawes, Z., Moss, J., Caswell, B., Naqvi, S., and MacKinnon, S. (2017). Enhancing children's spatial and numerical skills through a dynamic spatial approach to early geometry instruction: effects of a 32-week intervention. *Cogn. Instr.* 35, 236–264. doi: 10.1080/07370008.2017.1323902
- Hegarty, M., and Kozhevnikov, M. (1999). Types of visual-spatial representations and mathematical problem solving. *J. Educ. Psychol.* 91:684.
- Holloway, I. D., Price, G. R., and Ansari, D. (2010). Common and segregated neural pathways for the processing of symbolic and nonsymbolic numerical magnitude: an fMRI study. *Neuroimage* 49, 1006–1017. doi: 10.1016/j.neuroimage.2009.07.071
- Huttenlocher, J., and Presson, C. C. (1973). Mental rotation and the perspective problem. *Cogn. Psychol.* 4, 277–299. doi: 10.1016/0010-0285(73)90015-7
- Kaminski, J. A., Sloutsky, V. M., and Heckler, A. (2009). Transfer of mathematical knowledge: the portability of generic instantiations. *Child Dev. Perspect.* 3, 151–155. doi: 10.1111/j.1750-8606.2009.00096.x
- Kelley, T. D., Lee, F. J., and Wiley, P. (2000). *Developing an ACT-R Model of Spatial Manipulation* (ARL Tech. Rep. No. 2179). Aberdeen, MD: Aberdeen Proving Ground.
- Kline, W. E. (1960). *A synthesis of Two Factor Analyses of Intermediate Algebra*. Doctoral dissertation, Princeton University, University Microfilms, Ann Arbor.
- Kozhevnikov, M., and Hegarty, M. (2001). A dissociation between object manipulation spatial ability and spatial orientation ability. *Mem. Cogn.* 29, 745–756. doi: 10.3758/BF03200477
- Kozhevnikov, M., Kosslyn, S., and Shephard, J. (2005). Spatial versus object visualizers: a new characterization of visual cognitive style. *Mem. Cogn.* 33, 710–726. doi: 10.3758/BF03195337
- Kung, E., and Hamm, J. P. (2010). A model of rotated mirror/normal letter discriminations. *Mem. Cogn.* 38, 206–220. doi: 10.3758/MC.38.2.206
- Kyttälä, M., and Lehto, J. E. (2008). Some factors underlying mathematical performance: the role of visuospatial working memory and non-verbal intelligence. *Eur. J. Psychol. Educ.* 23:77. doi: 10.1007/BF03173141
- Kyttälä, M., Aunio, P., Lehto, J. E., Van Luit, J., and Hautamäki, J. (2003). Visuospatial working memory and early numeracy. *Educ. Child Psychol.* 20, 65–76.
- Landy, D., and Goldstone, R. L. (2010). Proximity and precedence in arithmetic. *Quart. J. Exp. Psychol.* 63, 1953–1968. doi: 10.1080/17470211003787619
- Levine, S. C., Goldin-Meadow, S., Carlson, M. T., and Hemani-Lopez, N. (2018). Mental transformation skill in young children: the role of concrete and abstract motor training. *Cogn. Sci.* doi: 10.1111/cogs.12603. [Epub ahead of print].
- Lohman, D. F. (1988). "Spatial abilities as traits, processes, and knowledge," in *Advances in the Psychology of Human Intelligence*, Vol. 4, ed R. J. Sternberg (Hillsdale, NJ: Lawrence Erlbaum Associates), 181–248.
- Lovett, A., and Schultheis, H. (2014). "Modeling spatial abstraction during mental rotation," in *Proceedings of the 36th Annual Meeting of the Cognitive Science Society* (Quebec City, QC), 886–891.
- Lowrie, T., Logan, T., and Ramful, A. (2017). Visuospatial training improves elementary students' mathematics performance. *Br. J. Educ. Psychol.* 87, 170–186. doi: 10.1111/bjep.12142
- Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L., and Ansari, D. (2014). Numerical predictors of arithmetic success in grades 1–6. *Dev. Sci.* 17, 714–726. doi: 10.1111/desc.12152
- MacCallum, R. C., Widaman, K. F., Zhang, S., and Hong, S. (1999). Sample size in factor analysis. *Psychol. Methods* 4:84. doi: 10.1037/1082-989X.4.1.84
- Matthews, P. G., and Hubbard, E. M. (2017). Making space for spatial proportions. *J. Learn. Disabil.* 50, 644–647. doi: 10.1177/0022219416679133
- McGee, M. G. (1979). Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychol. Bull.* 86:889. doi: 10.1037/0033-2909.86.5.889
- Michael, W. B., Guilford, J. P., Fruchter, B., and Zimmerman, W. S. (1957). The description of spatial-visualization abilities. *Educ. Psychol. Meas.* 17, 185–199. doi: 10.1177/001316445701700202
- Mix, K. S., and Cheng, Y. L. (2012). "The relation between space and math: developmental and educational implications," in *Advances in Child Development Behavior*, Vol. 42, ed J. B. Benson (New York, NY: Elsevier).
- Mix, K. S., Levine, S. C., Cheng, Y. L., Young, C. J., Hambrick, D. Z., and Konstantopoulos, S. (2017a). The latent structure of spatial skills and mathematics: a replication of the two-factor model. *J. Cogn. Dev.* 18, 465–492. doi: 10.1080/15248372.2017.1346658
- Mix, K. S., Levine, S. C., Cheng, Y. L., Young, C., Hambrick, D. Z., Ping, R., et al. (2016). Separate but correlated: the latent structure of space and mathematics across development. *J. Exp. Psychol. Gen.* 145, 1206–1227. doi: 10.1037/xge0000182
- Mix, K. S., Smith, L. B., Stockton, J. D., and Barterian, J. A. (2017b). Grounding the symbols for place value: do concrete models help? *J. Cogn. Dev.* 18, 129–151. doi: 10.1080/15248372.2016.1180296
- Mix, K. S. (2010). "Spatial tools for mathematical thought," in *The Spatial Foundations of Language and Cognition*, eds K. S. Mix, L. B. Smith and M. Gasser (New York, NY: Oxford University Press), 41–66.
- Newcombe, N. S., and Shipley, T. F. (2015). "Thinking about spatial thinking: new typology, new assessments," in *Studying Visual and Spatial Reasoning for Design Creativity*, ed J. S. Gero (New York, NY: Springer), 179–192.
- Newcombe, N. S. (2010). Picture this: Increasing math and science learning by improving spatial thinking. *Am. Educ.* 34, 29.
- Peters, M. (1995). *Revised Vandenberg & Kuse Mental Rotations Test: Forms MRT-A to MRT-D (Technical Report)*. Guelph, ON: University of Guelph, Department of Psychology.
- Piazza, M., Pinel, P., Le Bihan, D., and Dehaene, S. (2007). A magnitude code common to numerosities and number symbols in human intraparietal cortex. *Neuron* 53, 293–305. doi: 10.1016/j.neuron.2006.11.022
- Provost, A., and Heathcote, A. (2015). Titrating decision processes in the mental rotation task. *Psychol. Rev.* 122, 735–754. doi: 10.1037/a0039706
- Raghubar, K. P., Barnes, M. A., and Hecht, S. A. (2010). Working memory and mathematics: a review of developmental, individual difference, and cognitive approaches. *Learn. Individ. Differ.* 20, 110–122. doi: 10.1016/j.lindif.2009.10.005
- Reuhkala, M. (2001). Mathematical skills in ninth-graders: relationship with visuo-spatial abilities and working memory. *Educ. Psychol.* 21, 387–399.
- Richland, L. E., Stigler, J. W., and Holyoak, K. J. (2012). Teaching the conceptual structure of mathematics. *Educ. Psychol.* 47, 189–203. doi: 10.1080/00461520.2012.667065
- Rittle-Johnson, B., Fyfe, E. R., Hofer, K. G., and Farran, D. C. (2017). Early math trajectories: low-income children's mathematics knowledge from ages 4 to 11. *Child Dev.* 88, 1727–1742. doi: 10.1111/cdev.12662
- Rummel, R. J. (1970). *Applied factor Analysis*. Evanston, IL: Northwestern University Press.
- Rusch, C. E. (1957). *An Analysis of Arithmetic Achievement in Grades Four, Six, and Eight*. Doctoral dissertation, University of Wisconsin University Microfilms, Ann Arbor, 57–3568.

- Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., et al. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: a meta-analysis. *Dev. Sci.* 20:e12515. doi: 10.1111/desc.12372
- Schoenfeld, A. H. (1985). *Mathematical Problem Solving*. Orlando, FL: Academic Press.
- Searle, J. A., and Hamm, J. P. (2012). Individual differences in the mixture ratio of rotation and nonrotation trials during rotated mirror/normal letter discriminations. *Mem. Cogn.* 40, 594–613. doi: 10.3758/s13421-011-0172-2
- Sella, F., Sader, E., Lolliot, S., and Cohen Kadosh, R. (2016). Basic and advanced numerical performances relate to mathematical expertise but are fully mediated by visuospatial skills. *J. Exp. Psychol. Learn. Mem. Cogn.* 42, 1458–1472. doi: 10.1037/xlm0000249
- Shea, D. L., Lubinski, D., and Benbow, C. P. (2001). Importance of assessing spatial ability in intellectually talented young adolescents: a 20-year longitudinal study. *J. Educ. Psychol.* 93, 604–614. doi: 10.1037/0022-0663.93.3.604
- Shepard, R. N., and Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science* 171, 701–703. doi: 10.1126/science.171.3972.701
- Shepard, S., and Metzler, D. (1988). Mental rotation: effects of dimensionality of objects and type of task. *J. Exp. Psychol. Hum. Percept. Perform.* 14:3. doi: 10.1037/0096-1523.14.1.3
- Siegler, R. S., and Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychol. Sci.*, 14, 237–250. doi: 10.1111/1467-9280.02438
- Simons, D. J., Boot, W. R., Charness, N., Gathercole, S. E., Chabris, C. F., Hambrick, D. Z., et al. (2016). Do “brain-training” programs work?. *Psychol. Sci. Pub. Inter.* 17, 103–186. doi: 10.1177/1529100616661983
- Skemp, R. R. (1961). Reflective intelligence and mathematics. *Br. J. Educ. Psychol.* 31, 45–55. doi: 10.1111/j.2044-8279.1961.tb01701.x
- Slusser, E. B., Santiago, R. T., and Barth, H. C. (2013). Developmental change in numerical estimation. *J. Exp. Psychol. Gen.* 142, 193–208. doi: 10.1037/a0028560
- Sorby, S., Casey, B., Veurink, N., and Dulaney, A. (2013). The role of spatial training in improving spatial and calculus performance in engineering students. *Learn. Individ. Differ.* 26, 20–29. doi: 10.1016/j.lindif.2013.03.010
- Spearman, C. (1927). *The Abilities of Man*. Oxford: Macmillan.
- Stieff, M. (2013). Sex differences in the mental rotation of chemistry representations. *J. Chem. Educ.* 90, 165–170. doi: 10.1021/ed300499t
- Thompson, J. M., Nuerk, H. C., Moeller, K., and Kadosh, R. C. (2013). The link between mental rotation ability and basic numerical representations. *Acta Psychol.* 144, 324–331. doi: 10.1016/j.actpsy.2013.05.009
- Thurstone, L. L. (1938). *Primary Mental Abilities*. Chicago: University of Chicago Press.
- Tomarken, A. J., and Waller, N. G. (2005). Structural equation modeling: strengths, limitations, and misconceptions. *Annu. Rev. Clin. Psychol.* 1, 31–65. doi: 10.1146/annurev.clinpsy.1.102803.144239
- Tosto, M. G., Hanscombe, K. B., Haworth, C. M., Davis, O. S., Petrill, S. A., Dale, P. S., et al. (2014). Why do spatial abilities predict mathematical performance?. *Dev. Sci.* 17, 462–470. doi: 10.1111/desc.12138
- Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., et al. (2013). The malleability of spatial skills: a meta-analysis of training studies. *Psychol. Bull.* 139, 352–402. doi: 10.1037/a0028446
- Vandenberg, S. G., and Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Percept. Mot. Skills* 47, 599–604.
- Verdine, B. N., Irwin, C. M., Golinkoff, R. M., and Hirsh-Pasek, K. (2014). Contributions of executive function and spatial skills to preschool mathematics achievement. *J. Exp. Child Psychol.* 126, 37–51. doi: 10.1016/j.jecp.2014.02.012
- Wai, J., Lubinski, D., and Benbow, C. P. (2009). Spatial ability for STEM domains: aligning over 50 years of cumulative psychological knowledge solidifies its importance. *J. Educ. Psychol.* 101, 817–835. doi: 10.1037/a0016127
- Werdelin, I. (1966). *A Synthesis of Two Factor Analyses of Problem Solving in Mathematics*. Malmö: Malmö School of Education.
- Witt, M. (2011). School based working memory training: Preliminary finding of improvement in children's mathematical performance. *Adv. Cogn. Psychol.* 7, 7–15. doi: 10.2478/v10053-008-0083-3
- Xu, Y., and Franconeri, S. L. (2015). Capacity for visual features in mental rotation. *Psychol. Sci.* 26, 1241–1251. doi: 10.1177/0956797615585002
- Xu, C., and LeFevre, J. A. (2016). Training young children on sequential relations among numbers and spatial decomposition: differential transfer to number line and mental transformation tasks. *Dev. Psychol.* 52, 854–866. doi: 10.1037/dev0000124

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Young, Levine and Mix. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



A Taxonomy Proposal for Types of Interactions of Language and Place-Value Processing in Multi-Digit Numbers

Julia Bahnmüller^{1,2,3*}, Hans-Christoph Nuerk^{1,2,3} and Korbinian Moeller^{1,2,3}

¹ Neuro-cognitive Plasticity Laboratory, Leibniz-Institut für Wissensmedien, Tübingen, Germany, ² Department of Psychology, Eberhard Karls University of Tübingen, Tübingen, Germany, ³ LEAD Graduate School and Research Network, University of Tübingen, Tübingen, Germany

OPEN ACCESS

Edited by:

Julia Mary Carroll,
Coventry University, United Kingdom

Reviewed by:

Ann Dowker,
University of Oxford, United Kingdom
Olaf Hauk,
University of Cambridge,
United Kingdom

*Correspondence:

Julia Bahnmüller
j.bahnmüller@iwm-tuebingen.de

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 31 January 2018

Accepted: 31 May 2018

Published: 25 June 2018

Citation:

Bahnmüller J, Nuerk H-C and
Moeller K (2018) A Taxonomy
Proposal for Types of Interactions
of Language and Place-Value
Processing in Multi-Digit Numbers.
Front. Psychol. 9:1024.
doi: 10.3389/fpsyg.2018.01024

Research on associations between language and number processing has seen growing interest in the last years – in particular with respect to place-value processing in multi-digit numbers. Recently, Dowker and Nuerk (2016) proposed a taxonomy of linguistic influences on number processing. However, this taxonomy does not address the generality or specificity of linguistic influences across different levels of number processing. In contrast, Nuerk et al. (2015) proposed different levels of place-value processing in multi-digit numbers. However, the authors did not specify if and how linguistic factors influence these levels of place-value processing. The present perspective aims at addressing this conceptual gap by suggesting an integrated taxonomy representing how different linguistic factors may influence different levels of place-value processing. We show that some effects of different linguistic levels have already been observed on different levels of place-value processing. Moreover, while some linguistic influences (e.g., lexical influences) have been studied for all levels of place-value processing, other influences have been studied for only one level or even none. Beyond categorizing existing research, we argue that the explicit consideration of research gaps may inspire new research paradigms complementing the picture of language influences on place-value processing. We conclude by outlining the importance of a differential approach for levels of both linguistic and number processing to evaluate linguistic obstacles and facilitators of different languages and their relevance for numerical development.

Keywords: linguistic influences, numerical processing, place-value processing, multi-digit numbers, number word inversion

INTRODUCTION

Linguistic or language influences have seen growing research interest in the area of number processing and particularly with regard to place-value processing in multi-digit numbers. A systematic classification of levels of linguistic influences as well as their direction was recently proposed by Dowker and Nuerk (2016). However, the primary focus in numerical cognition research was (and still often is) on single-digit number processing. This may be problematic, because findings and conclusions obtained from research on single-digit numbers cannot simply be transferred to multi-digit numbers (cf. Nuerk et al., 2011). For instance, the majority of difficulties

in numerical development specifically relate to numbers and procedures beyond the single-digit number range (e.g., Zuber et al., 2009; for transcoding). For multi-digit Arabic numbers, one specifically crucial concept that needs to be acquired and understood is their place-value structuring principle. This principle reflects that the magnitude of a digit within the digit string (and consequently also of the overall number) can only be derived if spatial information regarding the position of digits within the digit string is considered. In particular, the spatial sequence of digits determines the value of a specific digit in descending powers of the base 10 from left to right (e.g., $4242 = \{4\} \times 10^3 + \{2\} \times 10^2 + \{4\} \times 10^1 + \{2\} \times 10^0$). Importantly, different levels of processing place-value information were specified (Nuerk et al., 2015). Thus, not only are there different linguistic levels affecting (multi-digit) number processing, but there are also different levels of place-value processing that can and should be distinguished.

Therefore, we argue that it is necessary to specify levels both of linguistic influences and place-value processing which are addressed in a specific paradigm to be able to distinguish and classify conceptually (dis)similar mechanisms underlying associations of language and place-value processing in multi-digit symbolic numbers. Such a classification comes with the opportunity to evaluate whether every linguistic influence is indeed relevant to each level of place-value processing and/or whether linguistic influences affect (only) specific levels of place-value processing. As a starting point, we suggest integrating the previously proposed taxonomy on linguistic factors influencing number processing by Dowker and Nuerk (2016) and the classification of different levels of place-value processing by Nuerk et al. (2015).

LINGUISTIC LEVELS INTERACTING WITH (MULTI-DIGIT) NUMBER PROCESSING

Large-scale cross-cultural studies, like TIMSS or PISA (e.g., Mullis et al., 2012; OECD, 2014) showed repeatedly that mathematical competences of children vary considerably between countries. One of the main and consistent findings is the superiority in mathematic performance of countries such as China, Japan, or Korea, also called the “Chinese number advantage” (e.g., Miura et al., 1993; Miura and Okamoto, 2003). Over and above educational systems and socio-economic factors (e.g., Towse and Saxton, 1998; Miller et al., 2005; Ngan Ng and Rao, 2010), linguistic specificities have been suggested to impact mathematical performance in general and place-value processing in particular. To specifically classify associations between linguistic specificities and number processing, Dowker and Nuerk (2016) recently introduced a taxonomy of six different linguistic levels: (A) lexical, (B) visuo-spatial orthographic (C) phonological (D) semantic (E) conceptual, and (F) syntactic.

The *lexical* level is the most widely investigated and is concerned with specificities on the number word level with respect to the transparency of power (e.g., in Chinese, power is explicit in number symbols and words: $42 = 四十二 = sì shí$

$èr = 4 \ 10 \ 2$) and transparency of order (e.g., the inversion of number words: in German the number word corresponding to 42 is *zweiundvierzig*, literally two and forty). The *visuo-spatial orthographic* level is not a typical linguistic category in most of the linguistic literature. However, this level includes effects of reading and writing direction and reading behavior that have been shown to heavily influence spatial-numerical processing (e.g., determining the direction of spatial numerical associations, Shaki et al., 2009). The *phonological* level summarizes effects of phonological processes and/or deficits as well as effects related to verbal working memory. Influences on the phonological level are, for instance, reflected by effects of concurrent articulation on specific aspects of number processing – indicating their reliance on verbal/phonological processing (e.g., Moeller et al., 2011). The *semantic* level is concerned with influences and characteristics of words (other than number words) and symbols that convey numerical meaning (e.g., more, less, buy, sell, +, –, cm, m). In this context, Shikhare et al. (2015) showed, for instance, that numerical estimation and comparison strategies as well as quantifier semantics determine the processing of proportional quantifiers (e.g., “few”, “many”, and “some”). Numerical processing is also influenced by certain linguistic *concepts* such as, for instance, linguistic markedness [e.g., there are unmarked (even, right) and marked forms (odd, left) of most adjective pairs]. Here, the effect of linguistic markedness of response codes (MARC effect, Nuerk et al., 2004) describes the finding that responses are faster for congruent pairings (i.e., even number/right hand response, odd number/left hand response) than incongruent ones (i.e., even number/left hand response, odd number/right hand response). Finally, the *syntactic* level refers to influences of grammar resulting from, for instance, specificities of certain grammatical rules. In this context, grammatical number was found to support learning cardinality of small numbers using the give-N task which requires the processing of the respective magnitude information. Sarnecka et al. (2007) compared groups of three-year-olds speaking languages with (English, Russian) and without plural markings (Japanese) and showed that more English/Russian than Japanese children gave the correct number of items indicating that grammar may have facilitated the acquisition of number cardinality.

In sum, the above taxonomy illustrates that language may be associated with number processing at different linguistic levels. As such, the term association is used intentionally to underline the potential bidirectionality of influences. Importantly, linguistic levels do not have to be task-relevant but might still influence the way we process numbers in a highly automatic yet implicit manner (as for reading direction and spatial-numerical associations; Shaki et al., 2009). Finally, more than one linguistic level may be associated with number processing: Moeller et al. (2015a) showed both lexical (inversion) as well as visuo-spatial orthographic (reading direction) influences on performance in a multi-digit number comparison task. However, place-value processing in multi-digit numbers is not unidimensional either. Both task requirements and processing characteristics that are specific to a respective task play a crucial role, and thus, linguistic levels might be equally important for some but not all tasks.

PLACE-VALUE PROCESSING LEVELS IN MULTI-DIGIT NUMBERS

Regarding multi-digit numbers, Nuerk et al. (2015) suggested three different levels of place-value processing to classify different tasks according to processing requirements: (1) place identification, (2) place-value activation, and (3) place-value computation.

Place identification is suggested to be an early and very basic requirement of virtually all tasks involving multi-digit numbers. This process is required for correctly identifying the position of a single digit within the digit string (e.g., tens and units positions in two-digit numbers) without the necessity of further processing the magnitude of these digits. An exemplary task involving place identification is transcoding of multi-digit numbers (i.e., writing numbers to dictation). With respect to transcoding, Nuerk and colleagues suggest that although magnitude information (by means of place-value activation) may be processed in addition to place identification, magnitude processing is not necessary (see also Cipolotti and Butterworth, 1995; Barrouillet et al., 2004).

In contrast to transcoding, other tasks such as number magnitude comparison require the *activation of place-value information*, which means that each symbol (digit) is associated with a specific position (place). Without place-value activation, the “Which number is larger?” question simply cannot be answered.

Finally, some tasks additionally require *place-value computation* in terms of changes or updates of value and/or place. For example, to correctly execute a carry operation in an addition task, the decade digit of the unit sum needs to be added to the sum of the decade digits to correctly solve the task (e.g., for $28+17$, $8+7=15$, and thus the sum of the decade digits needs to be updated accordingly, i.e., $2+1+1=4$). As such, carry problems are more difficult than non-carry problems (e.g., Deschuyteneer et al., 2005).

Taken together, there are different levels of linguistic influences on number processing and different levels of place-value processing for multi-digit numbers. Therefore, we suggest classifying any interaction of language and place-value processing in multi-digit numbers according to both, the level of linguistic influence and place-value processing.

INTEGRATING LEVELS OF LINGUISTIC INFLUENCES AND PLACE-VALUE PROCESSING

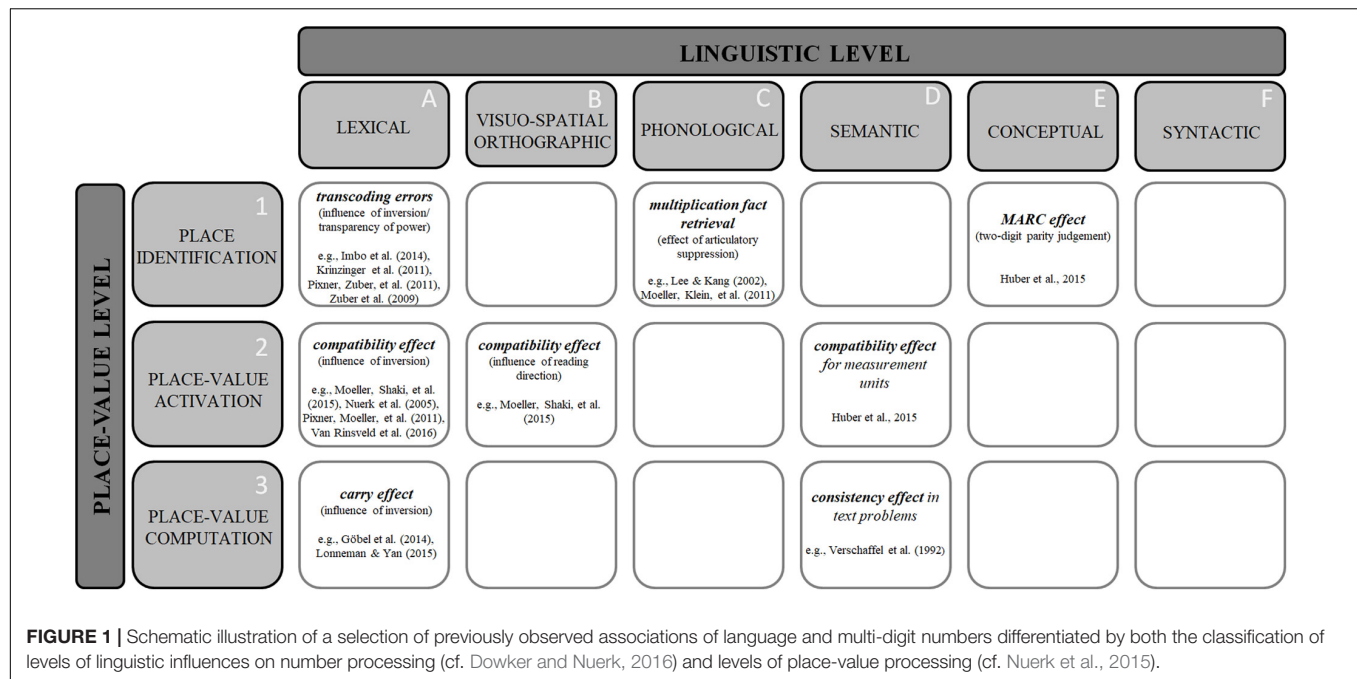
Classifying processes underlying different tasks and manipulations according to both linguistic and place-value processing levels results in a grid as depicted in **Figure 1**. Therein, each cell describes the association of one specific level of linguistic influences (A to F) with one specific level of place-value processing (1 to 3). It becomes evident that some associations have already been studied quite extensively, while others have been addressed only rarely or not at all so far.

A closer look at the studies investigating linguistic influences on (multi-digit) number processing indicates that two major approaches can be distinguished: first, cross-linguistic studies, comparing number processing effects across different languages or cultures, and second, linguistic manipulations that vary specific linguistic features within one language and evaluate differential effects on number processing.

Cross-Linguistic Approaches

Interestingly, the lexical and the visuo-spatial orthographic levels are dominated by cross-linguistic approaches focusing on number processing effects that are sensitive to influences of specific aspects of language systems. On the *lexical level*, two important aspects that vary between number word systems have been shown to influence place-value processing: transparency of power and transparency of order. Detrimental influences of nontransparent number word systems were identified in a variety of tasks and paradigms on all three levels of place-value processing. Regarding the association of place identification and lexical influences (**Figure 1, A1**), transcoding performance was shown to be specifically vulnerable to inversion-related errors (i.e., writing down 45 when dictated 54, e.g., Zuber et al., 2009; Krinzinger et al., 2011; Pixner et al., 2011b; Imbo et al., 2014; for specific errors in Japanese, see Moeller et al., 2015b). Moreover, with respect to place-value activation (**Figure 1, A2**), specific differences between inverted and non-inverted languages were observed for the unit-decade compatibility effect in two-digit number magnitude comparison [i.e., compatible number pairs (32_57 , $3 < 5$ and $2 < 7$) are responded to faster than incompatible pairs (37_62 , $3 < 6$ but $7 > 2$); Nuerk et al., 2001]. For both children (Pixner et al., 2011a) and adults (Nuerk et al., 2005; Moeller et al., 2015a; but see Ganor-Stern and Tzelgov, 2011) it was found that interference due to the irrelevant unit digit is more pronounced for languages with an inverted number word system. Finally, lexical influences were also investigated at the level of place-value computation (**Figure 1, A3**). For instance, Göbel et al. (2014) observed that the carry effect was more pronounced in German- (inverted number words) than Italian-speaking children (no inversion; see also Colomé et al., 2010; Lonnemann and Yan, 2015).

Investigations of influences of reading/writing direction on number processing (reflecting the *visual-spatial orthographic level*) have their origin in the assumption of a mental number line on which numbers are arranged from left to right in ascending order. This metaphor indicates a close association of numbers and space. Evidence for this claim comes, for example, from the SNARC effect (Dehaene et al., 1993), showing that in Western cultures smaller numbers are usually associated with the left-hand side, whereas larger numbers are associated with the right-hand side (for visual-spatial orthographic influences on spatial-numerical associations see Göbel et al., 2011; but see, e.g., van Dijck and Fias, 2011 for a working memory account and Schroeder et al., 2017 for a multiple coding account on the SNARC-effect). Regarding multi-digit numbers and with respect to the level of place-value activation, Moeller et al. (2015a) considered both visual-spatial orthographic (reading direction) and lexical influences (inversion) in a quadrilingual



cross-cultural study with German- and English-speaking adults (left-to-right reading languages with inverted and non-inverted number words, respectively) as well as Hebrew and Arabic speakers (right-to-left read languages with inverted and non-inverted number words, respectively; **Figure 1**, B2 and A2). Results indicated that compatibility effects were larger when the order of digits in symbolic Arabic notation did not match the order of tens and units in number words (i.e., German and Hebrew). Importantly, this study illustrates that levels of linguistic influences should not be considered in isolation because more than one linguistic level might actually impact number processing at the same time.

It is important to note that not every cross-linguistic study is also cross-cultural. First, samples can be chosen for which the cultural environment is held constant. For instance, Mark and Dowker (2015) investigated linguistic influences on mathematical development between language groups but within the same culture and educational system. In particular, Mark and Dowker (2015) compared children that spoke Chinese at home and learnt to count in Chinese at school to children that spoke Chinese at home and learnt to count in English at school. Therefore, major cultural discrepancies (e.g., educational system, cultural environment) were balanced between the two samples (for similar within-culture approaches see Dowker et al., 2008; Colomé et al., 2010; Pixner et al., 2011b; Imbo et al., 2014). Second, the investigation of bilingual speakers also allows for an investigation of cross-linguistic differences within one and the same culture (e.g., Macizo et al., 2010a,b; Macizo et al., 2011a,b; Van Rinsveld et al., 2016). Crucially, when investigating bilingual speakers not only differences between numerical processing in the respective languages but also potential cross-linguistic modulations can be evaluated [e.g., whether or not specificities of one language influence (numerical) processing in the other

language; cf. Van Rinsveld et al., 2016]. Such cross-linguistic modulations might have important implications for practical interventions for bilingual speakers. In general, research on cross-linguistic, though not cross-cultural studies substantiated influences of lexical linguistic properties on all three levels of place-value processing.

Language Manipulations

Instead of employing quasi-experimental designs comparing different language groups (or the same group in different language contexts) as described above, specific linguistic attributes may also be manipulated directly within one and the same language to identify additional interactions of linguistic and place-value processing levels. In particular, specific manipulations of phonological or semantic input as well as the consideration of specific linguistic concepts have already unraveled a variety of additional associations between levels of linguistic and place-value processing.

On the *phonological level*, for instance, Lee and Kang (2002) manipulated the availability of verbal information processing resources in multiplication and subtraction tasks and observed that concurrent articulation specifically reduced multiplication fact retrieval but not subtraction performance. This indicates that phonological processing of number words indeed affects place-value processing in multi-digit numbers differentially and even when no explicit magnitude processing is required to correctly solve the task (see also Moeller et al., 2011; **Figure 1**, C3).

Next to insights resulting from the manipulation of phonological processing resources, interactions of levels of linguistic influences and place-value processing were explored by considering stimuli that are semantically different from Arabic numbers or number words but still convey numerical meaning. By manipulating the semantic input, investigations on the

semantic level allow for both an identification of effects resulting from specific word categories and/or for a generalization of number processing effects across different words/symbols. Referring to the former, for text problems it was observed that words associated with an addition procedure (e.g., “more,” “buy”) facilitated processing of text problems requiring additions whereas words associated with subtraction (e.g., “less,” “sell”) interfered with addition problem solving (e.g., Verschaffel et al., 1992; see also Daroczy et al., 2015 for a review on linguistic and numerical factors in text problems; **Figure 1**, D3). Referring to the latter, place-value processing also seems to be recycled for the processing of measurement units as typical effects observed for two-digit numbers (e.g., unit-decade compatibility effect) were also demonstrated for measurement units (Huber et al., 2015a; **Figure 1**, D2). Thus, these studies show that magnitude information is not only expressed and processed via Arabic digits and number words but also via other words and symbols which in some cases share processing specificities observed for place-value processing in multi-digit numbers.

Finally, there is first evidence for an interaction between linguistic aspects and place-identification on the *conceptual level*, specifically through manipulating the markedness of response codes. Huber et al. (2015b) investigated the MARC effect in a two-digit parity judgement task. A regular MARC effect was observed for both single- and (the unit digit of) two-digit numbers (**Figure 1**, E1). This suggests that the manipulation of specific linguistic concepts might interfere with place-value identification as well.

FILLING THE GAPS: INSPIRING FUTURE RESEARCH

In addition to assigning different levels of linguistic and place-value processing to categorize existing research, a taxonomy may also inspire future research. For instance, addressing gaps at the visuo-spatial orthographic level, cross-cultural studies using a quadrilingual design comparable to the one used in Moeller et al. (2015a) might help to evaluate questions on the generality of visual-spatial orthographic influences across different place-value processing levels. Using, for instance, transcoding in children and/or addition tasks should allow for investigating influences on place-identification and the place-value manipulation level, respectively (i.e., **Figure 1**, B1 and B3).

Moreover, future research might also consider combining not only different linguistic levels but also different approaches (e.g., combining quasi-experimental and experimental designs). For example, it would be interesting to evaluate whether a linguistic effect determined on one linguistic level and in one language group generalizes to or differs from other language groups. On the syntactic level, for instance, effects of specific grammatical structures were found to influence processing of single-digit numbers (e.g., Sarnecka et al., 2007). However, these effects have not yet been investigated for multi-digit numbers. Potential syntactic effects on place-value processing might be investigated in language groups with differing ways of expressing grammatical number. For instance, in many languages, the singular is used in

relation to one entity and plural for entities larger than one. In contrast, in Polish, the unit digits 2 to 4 are followed by plural verb forms whereas for the unit digits 1 and 5 to 9 singular is used. The same pattern holds for multi-digit numbers with the respective unit digits (e.g., 22 to 24 is followed by plural verb forms; 21 and 25 to 29 are followed by singular verb forms). In this context, the grammatical SNARC effect (i.e., singular associated with left and plural with right; Roettger and Domahs, 2015) might be investigated in a cross-linguistic study design to determine the language specificity of this syntactic effect and its potential generalizability to the multi-digit number range.

Finally, next to a broadening of our understanding of the generality and limits of interactions of linguistic and place-value processing levels, it will also be crucial to identify developmental trajectories of such interactions as well as their different effect sizes, i.e. their differential significance in practical contexts to be able to develop tailored types and time windows for potential interventions.

CONCLUSION

Language considerably influences numerical cognition and development. Therefore, we suggest that it is important to understand the principles of such influences in any language. To foster such understanding, the goal of this article was to show that the general conceptualization, “language influences multi-digit number processing” captures neither the diversity of different levels of linguistic influences nor that of different levels of place-value processing. So far, a lot of research effort has been devoted to investigating prominent linguistic influences (mostly lexical), and has to a large part neglected others. We hope that this overview and taxonomy inspires researchers to study other linguistic influences on different levels of place-value processing as well to generate a more complete and differentiated picture of such interactions in the future. This will help us to better understand benefits and obstacles for numerical and arithmetic processing and learning in a given language and ultimately foster development and remediation tailored to each language background as well.

AUTHOR CONTRIBUTIONS

All authors contributed intellectually to the conceptualization and revision of this perspective paper and read and approved the submitted version. The initial draft was written by JB.

FUNDING

JB was supported by the Leibniz-Competition Fund (SAW-2014IWM-4) providing funding to Elise Klein. H-CN and KM were principal investigators and JB associated member at the LEAD Graduate School and Research Network [GSC1028], a project of the Excellence Initiative of the German Federal and State Governments.

REFERENCES

- Baddeley, A., Lewis, V., and Vallar, G. (1984). Exploring the articulatory loop. *Q. J. Exp. Psychol. A* 36, 233–252. doi: 10.1080/14640748408402157
- Barrouillet, P., Camos, V., Perruchet, P., and Seron, X. (2004). ADAPT: a developmental, asemantic, and procedural model for transcoding from verbal to Arabic numerals. *Psychol. Rev.* 111, 368–394. doi: 10.1037/0033-295X.111.2.368
- Cipolotti, L., and Butterworth, B. (1995). Toward a multiroute model of number processing: impaired number transcoding with preserved calculation skills. *J. Exp. Psychol. Gen.* 124, 375–390. doi: 10.1037/0096-3445.124.4.375
- Colomé, À., Laka, I., and Sebastián-Gallés, N. (2010). Language effects in addition: how you say it counts. *Q. J. Exp. Psychol.* 63, 965–983. doi: 10.1080/17470210903134377
- Daroczy, G., Wolska, M., Meurers, W. D., and Nuerk, H.-C. (2015). Word problems: a review of linguistic and numerical factors contributing to their difficulty. *Front. Psychol.* 6:348. doi: 10.3389/fpsyg.2015.00348
- Dehaene, S., Bossini, S., and Giraux, P. (1993). The mental representation of parity and number magnitude. *J. Exp. Psychol. Gen.* 122, 371–396. doi: 10.1037/0096-3445.122.3.371
- Deschuyteneer, M., De Rammelaere, S., and Fias, W. (2005). The addition of two-digit numbers: exploring carry versus no-carry problems. *Psychol. Sci.* 47, 74–83.
- Dowker, A., Bala, S., and Lloyd, D. (2008). Linguistic influences on mathematical development: how important is the transparency of the counting system? *Philos. Psychol.* 21, 523–538. doi: 10.1080/09515080802285511
- Dowker, A., and Nuerk, H.-C. (2016). Editorial: linguistic influences on mathematics. *Front. Psychol.* 7:1035. doi: 10.3389/fpsyg.2016.01035
- Ganor-Stern, D., and Tzelgov, J. (2011). Across-notation automatic processing of two-digit numbers. *Exp. Psychol.* 58, 147–153. doi: 10.1027/1618-3169/a000080
- Göbel, S. M., Moeller, K., Pixner, S., Kaufmann, L., and Nuerk, H.-C. (2014). Language affects symbolic arithmetic in children: the case of number word inversion. *J. Exp. Child Psychol.* 119, 17–25. doi: 10.1016/j.jecp.2013.10.001
- Göbel, S. M., Shaki, S., and Fischer, M. H. (2011). The cultural number line: a review of cultural and linguistic influences on the development of number processing. *J. Cross Cult. Psychol.* 42, 543–565. doi: 10.1177/0022022111406251
- Huber, S., Bahnmüller, J., Klein, E., and Moeller, K. (2015a). Testing a model of componential processing of multi-symbol numbers - Evidence from measurement units. *Psychon. Bull. Rev.* 22, 1417–1423. doi: 10.3758/s13423-015-0805-8
- Huber, S., Klein, E., Graf, M., Nuerk, H.-C., Moeller, K., and Willmes, K. (2015b). Embodied markedness of parity? Examining handedness effects on parity judgments. *Psychol. Res.* 79, 963–977. doi: 10.1007/s00426-014-0626-9
- Imbo, I., Vanden Bulcke, C., De Brauwier, J., and Fias, W. (2014). Sixty-four or four-and-sixty? The influence of language and working memory on children's number transcoding. *Front. Psychol.* 5:313. doi: 10.3389/fpsyg.2014.00313
- Krinzinger, H., Gregoire, J., Desoete, A., Kaufmann, L., Nuerk, H.-C., and Willmes, K. (2011). Differential language effects on numerical skills in second grade. *J. Cross Cult. Psychol.* 42, 614–629. doi: 10.1177/0022022111406252
- Lee, K.-M., and Kang, S. (2002). Arithmetic operation and working memory: differential suppression in dual tasks. *Cognition* 83, B63–B68. doi: 10.1016/S00100277(02)00010-0
- Lonnemann, J., and Yan, S. (2015). Does number word inversion affect arithmetic processes in adults? *Trends Neurosci. Educ.* 4, 1–5. doi: 10.1016/j.tine.2015.01.002
- Macizo, P., Herrera, A., Paolieri, D., and Román, P. (2010a). Is there cross-language modulation when bilinguals process number words? *Appl. Psycholinguist.* 31, 651–669. doi: 10.1017/S0142716410000184
- Macizo, P., Herrera, A., Román, P., and Martín, M. C. (2010b). Second language acquisition influences the processing of number words. *Procedia Soc. Behav. Sci.* 9, 1128–1134. doi: 10.1016/j.sbspro.2010.12.295
- Macizo, P., Herrera, A., Román, P., and Martín, M. C. (2011a). Proficiency in a second language influences the processing of number words. *J. Cogn. Psychol.* 23, 915–921. doi: 10.1080/20445911.2011.586626
- Macizo, P., Herrera, A., Román, P., and Martín, M. C. (2011b). The processing of two-digit numbers in bilinguals. *Br. J. Psychol.* 102, 464–477. doi: 10.1111/j.2044-8295.2010.02005.x
- Mark, W., and Dowker, A. (2015). Linguistic influence on mathematical development is specific rather than pervasive: revisiting the Chinese Number Advantage in Chinese and English children. *Front. Psychol.* 6:203. doi: 10.3389/fpsyg.2015.00203
- Miller, K. F., Kelly, M., and Zhou, X. (2005). “Learning mathematics in China and the United States: cross-cultural insights into the nature and course of preschool mathematical development,” in *Handbook of Mathematical Cognition*, ed. J. I. D. Campbell (New York, NY: Psychology Press), 163–178.
- Miura, I. T., and Okamoto, Y. (2003). “Language supports for mathematics understanding and performance,” in *The Development of Arithmetic Concepts and Skills: Constructing Adaptive Expertise. Studies in Mathematics Thinking and Learning*, eds A. J. Baroody and A. Dowker (Mahwah, NJ: Erlbaum), 229–242.
- Miura, I. T., Okamoto, Y., Kim, C. C., Steere, M., and Fayol, M. (1993). First graders' cognitive representation of number and understanding of place value: cross-national comparisons: France, Japan, Korea, Sweden, and the United States. *J. Educ. Psychol.* 85, 24–30. doi: 10.1037/0022-0663.85.1.24
- Moeller, K., Klein, E., Fischer, M. H., Nuerk, H., and Willmes, K. (2011). Representation of multiplication facts - Evidence for partial verbal coding. *Behav. Brain Funct.* 7:25. doi: 10.1186/1744-9081-7-25
- Moeller, K., Shaki, S., Göbel, S. M., and Nuerk, H.-C. (2015a). Language influences number processing – A quadrilingual study. *Cognition* 136, 150–155. doi: 10.1016/j.cognition.2014.11.003
- Moeller, K., Zuber, J., Olsen, N., Nuerk, H.-C., and Willmes, K. (2015b). Intransparent German number words complicate transcoding - A translational comparison with Japanese. *Front. Psychol.* 6:740. doi: 10.3389/fpsyg.2015.00740
- Mullis, I. V. S., Martin, M. O., Foy, P., and Arora, A. (2012). *TIMSS 2011 International Results in Mathematics*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center.
- Ngan Ng, S. S., and Rao, N. (2010). Chinese number words, culture, and mathematics learning. *Rev. Educ. Res.* 80, 180–206. doi: 10.3102/0034654310364764
- Nuerk, H.-C., Iversen, W., and Willmes, K. (2004). Notational modulation of the SNARC and the MARC (linguistic markedness of response codes) effect. *Q. J. Exp. Psychol. A* 57, 835–863. doi: 10.1080/02724980343000512
- Nuerk, H.-C., Moeller, K., Klein, E., Willmes, K., and Fischer, M. H. (2011). Extending the mental number line - A review of multi-digit number processing. *Z. Psychol.* 219, 3–22. doi: 10.1027/2151-2604/a000041
- Nuerk, H.-C., Moeller, K., and Willmes, K. (2015). “Multi-digit number processing: overview, conceptual clarifications, and language influences,” in *The Oxford Handbook of Numerical Cognition*, eds R. C. Kadosh and A. Dowker (Oxford: Oxford University Press), 106–139.
- Nuerk, H.-C., Weger, U., and Willmes, K. (2001). Decade breaks in the mental number line? Putting the tens and units back in different bins. *Cognition* 82, B25–B33. doi: 10.1016/S0010-0277(01)00142-1
- Nuerk, H.-C., Weger, U., and Willmes, K. (2005). Language effects in magnitude comparison: small, but not irrelevant. *Brain Lang.* 92, 262–277. doi: 10.1016/j.bandl.2004.06.107
- OECD. (2014). *PISA 2012 Results: What Students Know and Can Do – Student Performance in Mathematics, Reading and Science*, Vol. I. Paris: PISA, OECD Publishing. doi: 10.1787/9789264201118-en
- Pixner, S., Moeller, K., Heřmanová, V., Nuerk, H.-C., and Kaufmann, L. (2011a). Language effects on nonverbal number processing in first grade - A trilingual study. *J. Exp. Child Psychol.* 108, 371–382. doi: 10.1016/j.jecp.2010.09.002
- Pixner, S., Zuber, J., Heřmanová, V., Kaufmann, L., Nuerk, H.-C., and Moeller, K. (2011b). One language, two number-word systems and many problems: numerical cognition in the Czech language. *Res. Dev. Disabil.* 32, 2683–2689. doi: 10.1016/j.ridd.2011.06.004
- Roettger, T. B., and Domahs, F. (2015). Grammatical number elicits SNARC and MARC effects as a function of task demands. *Q. J. Exp. Psychol.* 68, 1231–1248. doi: 10.1080/17470218.2014.979843
- Sarnecka, B. W., Kamenskaya, V. G., Yamana, Y., Ogura, T., and Yudovina, Y. B. (2007). From grammatical number to exact numbers: early meanings of ‘one’, ‘two’, and ‘three’ in English, Russian, and Japanese. *Cogn. Psychol.* 55, 136–168. doi: 10.1016/j.cogpsych.2006.09.001
- Schroeder, P. A., Nuerk, H.-C., and Plewnia, C. (2017). Switching between multiple codes of SNARC-like associations: two conceptual replication attempts with

- anodal tDCS in sham-controlled cross-over design. *Front. Neurosci.* 11:654. doi: 10.3389/fnins.2017.00654
- Shaki, S., Fischer, M. H., and Petrusic, W. M. (2009). Reading habits for both words and numbers contribute to the SNARC effect. *Psychon. Bull. Rev.* 16, 328–331. doi: 10.3758/PBR.16.2.328
- Shikhare, S., Heim, S., Klein, E., Huber, S., and Willmes, K. (2015). Processing of numerical and proportional quantifiers. *Cogn. Sci.* 39, 1504–1536. doi: 10.1111/cogs.12219
- Towse, J. N., and Saxton, M. (1998). “Mathematics across national boundaries: cultural and linguistic perspectives on numerical competence,” in *The Development of Mathematics Skills*, ed. C. Donlan (Hove: Psychology Press), 129–150.
- van Dijck, J. P., and Fias, W. (2011). A working memory account for spatial-numerical associations. *Cognition* 119, 114–119. doi: 10.1016/j.cognition.2010.12.013
- Van Rinsveld, A., Schiltz, C., Landerl, K., Brunner, M., and Ugen, S. (2016). Speaking two languages with different number naming systems: what implications for magnitude judgments in bilinguals at different stages of language acquisition? *Cogn. Process.* 17, 225–241. doi: 10.1007/s10339-016-0762-9
- Verschaffel, L., De Corte, E., and Pauwels, A. (1992). Solving compare problems: an eye movement test of Lewis and Mayer’s consistency hypothesis. *J. Educ. Psychol.* 84, 85–94. doi: 10.1037/0022-0663.84.1.85
- Zuber, J., Pixner, S., Moeller, K., and Nuerk, H.-C. (2009). On the language specificity of basic number processing: transcoding in a language with inversion and its relation to working memory capacity. *J. Exp. Child Psychol.* 102, 60–77. doi: 10.1016/j.jecp.2008.04.003

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Bahnmüller, Nuerk and Moeller. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



A Mental Odd-Even Continuum Account: Some Numbers May Be “More Odd” Than Others and Some Numbers May Be “More Even” Than Others

Lia Heubner^{1†}, Krzysztof Cipora^{1,2*†}, Mojtaba Soltanlou^{1,2}, Marie-Lene Schlenker¹, Katarzyna Lipowska³, Silke M. Göbel⁴, Frank Domahs⁵, Maciej Haman³ and Hans-Christoph Nuerk^{1,2,6}

¹ Department of Psychology, University of Tuebingen, Tuebingen, Germany, ² LEAD Graduate School and Research Network, University of Tuebingen, Tuebingen, Germany, ³ Department of Psychology, University of Warsaw, Warsaw, Poland, ⁴ Department of Psychology, University of York, York, United Kingdom, ⁵ Institute for German Linguistics, University of Marburg, Marburg, Germany, ⁶ Leibniz-Institut für Wissensmedien, Tuebingen, Germany

OPEN ACCESS

Edited by:

Ann Dowker,
University of Oxford, United Kingdom

Reviewed by:

Matthias Hartmann,
Universität Potsdam, Germany
Koen Luwel,
KU Leuven, Belgium

*Correspondence:

Krzysztof Cipora
krzysztof.cipora@uni-tuebingen.de

[†] These authors have contributed
equally to this work.

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 15 February 2018

Accepted: 07 June 2018

Published: 28 June 2018

Citation:

Heubner L, Cipora K, Soltanlou M, Schlenker M-L, Lipowska K, Göbel SM, Domahs F, Haman M and Nuerk H-C (2018) A Mental Odd-Even Continuum Account: Some Numbers May Be “More Odd” Than Others and Some Numbers May Be “More Even” Than Others. *Front. Psychol.* 9:1081. doi: 10.3389/fpsyg.2018.01081

Numerical categories such as parity, i.e., being odd or even, have frequently been shown to influence how particular numbers are processed. Mathematically, number parity is defined categorically. So far, cognitive, and psychological accounts have followed the mathematical definition and defined parity as a categorical psychological representation as well. In this manuscript, we wish to test the alternative account that cognitively, parity is represented in a more gradual manner such that some numbers are represented as “more odd” or “more even” than other odd or even numbers, respectively. Specifically, parity processing might be influenced by more specific properties such as whether a number is a prime, a square number, a power of 2, part of a multiplication table, divisible by 4 or by 5, and many others. We suggest that these properties can influence the psychologically represented parity of a number, making it more or less prototypical for odd- or evenness. In the present study, we tested the influence of these numerical properties in a bimanual parity judgment task with auditorily presented two-digit numbers. Additionally, we further investigated the interaction of these numerical properties with linguistic factors in three language groups (English, German, and Polish). Results show significant effects on reaction times of the congruity of parity status between decade and unit digits, even if numerical magnitude and word frequency are controlled. We also observed other effects of the above specific numerical properties, such as multiplication attributes, which facilitated or interfered with the speed of parity judgment. Based on these effects of specific numerical properties we proposed and elaborated a parity continuum account. However, our cross-lingual study also suggests that parity representation and/or access seem to depend on the linguistic properties of the respective language or education and culture. Overall, the results suggest that the “perceived” parity is not the same as objective parity, and some numbers are more prototypical exemplars of their categories.

Keywords: parity judgment, markedness, numerical properties, prototypicality, cross-linguistic comparisons

INTRODUCTION

Parity judgment—that is, deciding whether a number is even or odd—is one of the earliest mathematical tasks learned in school. Formally, parity can take one of two values: an even number is an integer of the form $n = 2k$, while an odd number is an integer of the form $n = 2k + 1$. Going further, in group theory¹, even and odd numbers build a ring with a zero element (neutral element of addition, i.e., even numbers) and a 1-element (neutral element of multiplication, i.e., odd numbers).

Thus, mathematically, parity is clearly defined. The aim of the present study was, however, to explore how parity is processed cognitively. While cognitive and psychological accounts so far have followed the mathematical definition and defined parity in terms of a categorical psychological representation, the present study aimed at testing an alternative account: Cognitively, parity may be represented in a more gradual manner, such that some numbers are represented as “more odd” or “more even” than other odd or even numbers, respectively. While this may seem an irritating concept for some numerical cognition researchers at first sight, we actually borrow from old ideas, which we apply to the concept of parity. Prototype theory (e.g., Posner and Keele, 1968; Rosch et al., 1976; Osherson and Smith, 1981) has long suggested that certain members of distinct categories are more typical examples of that category than others and that membership to such a category may be graded. Using such a theoretical conceptualization, a difference between formal binary categories and graded psychological processing can even be found in number processing, namely in processing numerical magnitude: The time needed to make (binary) same-different numerical judgments depends on the difference in magnitude between numbers (Dehaene and Akhavein, 1995, see also Sasanguie et al., 2011). Similarly, the time needed for numerical comparisons increases with decreasing distance between the numbers to be compared (numerical distance effect; Moyer and Landauer, 1967). However, for parity processing, such a graded account has—to the best of our knowledge—not been systematically tested yet (but see Armstrong et al., 1983 for an early account).

The Odd-Even Continuum: Tentative Account of the Influence of Numerical Properties on Perceived Parity Based on Prototypicality

Several studies conducted to date have suggested that participants' responses to the parity of different numbers vary. Smallest Space Analyses (SSA-I; Guttman, 1968; Lingoes and Roskam, 1973) conducted by Nuerk et al. (2004) show that zero is located further away (i.e., processed differently) from other numbers in a parity judgment task. While Nuerk et al. (2004) only suggested that the number zero is distinct, we wish to go beyond this claim here: We suggest that many more or maybe all numbers are represented differently with regard to parity on

a graded, continuous dimension. Indeed, as a small side claim in their seminal SNARC article, Dehaene et al. (1993) proposed that the mental representation of parity is influenced by several semantic properties, and pointed out that some numbers might be more prototypically odd or even. By extending this claim, one might hypothesize that specific properties facilitate or impede number processing, implying further that numbers are represented on an “oddness” or “evenness” continuum.

Dehaene et al. (1993) propose that prototypical numbers (i.e., numbers sharing many of the properties contributing to perceived parity) are classified faster as odd or even. One can postulate that one of the main factors contributing to perceived oddness and evenness would be the subjective ease of divisibility, as the parity concept itself strictly refers to divisibility by 2. The easier the division of a given number, the less subjectively odd/more subjectively even the number should be. This assumption meshes well with research on prototypicality (e.g., Rosch, 1975; Rosch and Lloyd, 1978) showing that some objects within a given category are categorized faster than others because they are (proto) typical exemplars of that category. To illustrate the point, among single-digit even numbers, 4, and 8 are powers of 2, potentially making them especially subjectively even. Only the number 6 in this set is not a power of 2 and is not divisible by 4, and as reported by Dehaene et al. (1993), the number 6 was an outlier in a parity judgment task, invoking exceptionally long reaction times. More recent studies have found that zero (Nuerk et al., 2004), 2², and 6 (reanalysis of data reported in Cipora and Nuerk, 2013) among even numbers are outliers prompting longer reaction times.

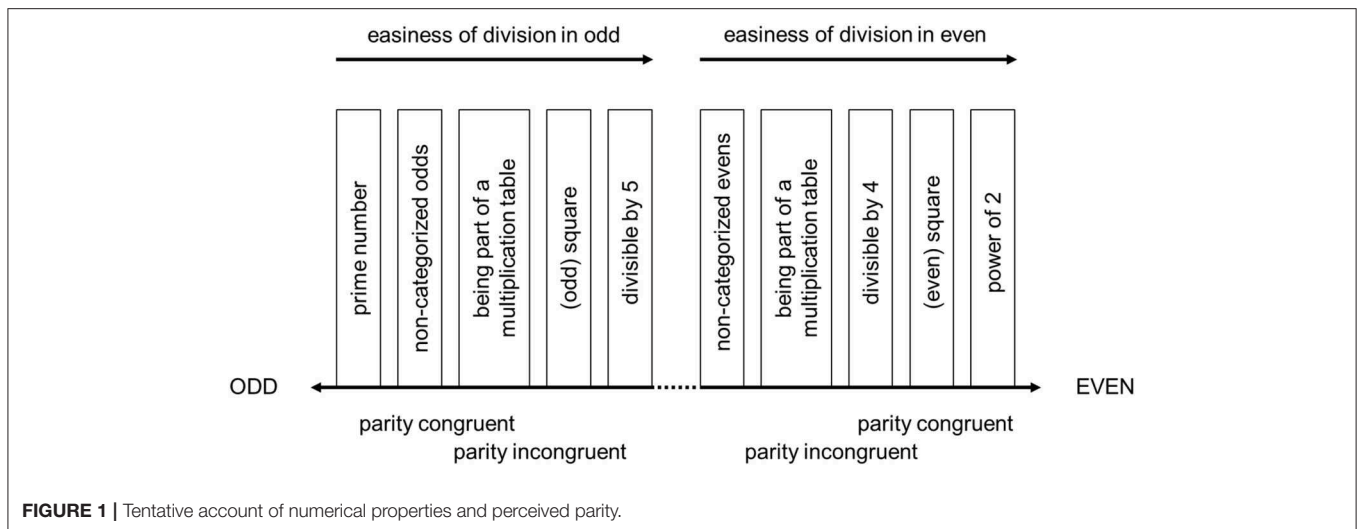
While some properties are expected to influence the perceived “evenness” of a number, other properties should influence the perceived “oddness.” For instance, whether a number is prime may contribute to its subjective oddness. Notably, numbers 1 and 9 are the only one-digit odd numbers that are not prime numbers, and a reanalysis of data reported by Cipora and Nuerk (2013) showed that in the case of odd numbers, reactions to the number 9 were the slowest among odd numbers. Dehaene et al. (1993) presented similar findings, with numbers 1 and 9 invoking longer reaction times than 3, 5 and 7.

These factors may explain the general patterns in one-digit numbers, but of course cannot be systematically tested in one-digit numbers, given that there are too few numbers and too many degrees of freedom (e.g., almost all one-digit odd numbers are also primes, almost all even one-digit numbers are also powers of two; see above). These confounds are also reflected in the inconclusive results of experiments using single-digit numbers. However, such assumptions can be tested for two-digit numbers, which we therefore set out to investigate here.

We suggest the “parity continuum” as a tentative account of the influence of numerical properties on the parity representation of two-digit numbers. In line with the properties investigated by

¹A group theory in mathematics is about understanding algebraic structures known as groups, which consist of a set of elements and an operation. Here, it provides a background and formal framework for a concept of parity.

²On the one hand this result is surprising as number two can be considered as a prototypical even number. On the other hand, it is also a prime number (i.e., it is divisible only by one and by itself). Even more importantly, number 2 is the only even prime number. This property may, at least in some individuals, lead to longer parity decision times to this number.



Dehaene et al. (1993), we included being a *prime number* (being divisible only by one and by itself, e.g., 23) and being a *power of 2* (e.g., 32, 64) as prototypical numerical properties for being odd and even, respectively. These two properties constitute extremes of perceived easiness of division (cf. **Figure 1**). Nevertheless, there are several other properties that can conceivably affect parity judgments, and which also influence the ease of division. These properties will be outlined in the following paragraphs.

With respect to easiness of division, it is easy to recognize numbers *divisible by 5* using a very simple heuristic. Furthermore, studies investigating the relationship between finger-counting habits and number processing suggest a key role of 5 as a sub-base in mental quantity representation and arithmetic. Such sub-base-5 effects have been observed in a number comparison task (Domahs et al., 2010), and a completing-addition production task (Klein et al., 2011). For these two reasons, we postulate that divisibility by 5 decreases the perceived oddness of a given number. At this stage in our tentative model we do not consider even numbers which are divisible by 5, that is full decades. In the base-10 system, decade numbers are special for many reasons (e.g., length, role in the base-10-system, e.g., for carry overs, consistency effects in multiplication and so forth; see e.g., Nuerk et al., 2002, 2015).

For even numbers, *divisibility by 4* also makes division more accessible, because the result of the division by 2 is also even (i.e., divisible by 2). This claim is supported by unpublished data collected by one of the co-authors (H-CN), which show that numbers divisible by 4 have unique characteristics compared to other even numbers. In that sense, divisibility by 4 increases perceived evenness of the number.

Following the easiness of division account, it must be noted that numbers that are *part of multiplication tables* are divisible by definition. Furthermore, in many educational systems, multiplication tables are learned by rote memorization. Therefore, we are more familiar with these numbers. They can be processed more easily than numbers we rarely encounter, which was shown in several studies. For instance, in number bisection tasks, participants tend to respond faster and more

accurately to items with numbers that are part of the same multiplication table (Nuerk et al., 2002). Therefore, being part of the multiplication table decreases perceived oddness and increases the perceived evenness of the number. Furthermore, even numbers constitute the majority (75%) of results of the multiplication table (because odd \times odd is the only combination leading to an odd multiplication result). In line with the easiness of division and familiarity notions, we also added being a *square number* to the account. As French (2005) points out, special attention is put on square numbers in mathematics lessons, which increases their familiarity and, akin to the other numbers that are part of a multiplication table, might influence their prototypicality and, thus, how their parity is processed. Being a square may decrease the perceived oddness and increase the perceived evenness of a number, because even numbers are probably generally more familiar. In **Figure 1** we present a tentative model of the parity continuum account, in which apart from the abovementioned properties we included the postulated positions of odd and even numbers that are not characterized by any of them. The order of categories depends on the postulated easiness of division within both odd and even numbers.

Empirical studies on parity judgments in two-digit numbers indicate that more than just the mathematical properties of the number influence reaction times. Namely, participants tend to respond faster to two-digit numbers if the number's decade and unit have the same parity status (both even: e.g., 48; both odd: e.g., 73), and respond slower if the parity status of the decade and the unit differ from each other (one even, one odd: e.g., 32, 45; Dehaene et al., 1993; Tan and Dixon, 2011). This effect, referred to as *parity congruity* is one of the 17 effects suggested to indicate decomposed processing of multi-digit numbers (Nuerk et al., 2011a,b for reviews). Although it is not an attribute related to division and multiplication (and therefore not depicted in **Figure 1**), parity congruity influences the ease of the parity decision and needs to be taken into account.

To sum up, properties related to divisibility, sub-base and familiarity as well as parity congruity seem to influence the perceived parity of two-digit numbers. What is more, one can

point to a number of linguistic factors that need to be taken into consideration while investigating numerical processing.

Linguistic Factors Influencing Number Processing

Numerical processing is also affected by linguistic features (see e.g., Dowker and Nuerk, 2016). Previous studies show that processing (i.e., accessing and operating with) even numbers might be different from processing odd numbers. One of the effects explained by linguistic factors is the so called “odd effect”: In a traditional parity judgment task, people tend to respond faster to even numbers than to odd numbers (Hines, 1990). It is often explained by the concept of linguistic markedness. It assumes that adjectives are arranged in pairs, which contain a marked, basic form and an unmarked one—the derived one. The unmarked form is the “more natural” form of an adjective and the marked form in some cases can be even produced out of the unmarked form by adding a negation prefix. In other cases, the marked form is identified as being less frequent (e.g., we ask “How old are you?”/“How long does it take?” rather than “How young are you?”/“How short does it take?” see e.g., Nuerk et al., 2004; Huber et al., 2015; Schroeder et al., 2017). It is also relatively easy to indicate markedness of adjectives referring to number parity. Evenness is considered as the unmarked form, and oddness as a marked one. In English, the word odd apart from denoting numbers indivisible by 2, means also “weird” or “non-typical”). In German and in Polish, adjectives denoting odd numbers are built by adding negation prefixes to adjectives denoting even numbers (“*ungerade*” and “*nieparzysty*” respectively). As shown in previous studies, the unmarked adjective-forms can be retrieved faster (Sherman, 1976), possibly explaining why even (“unmarked”) numbers are responded to faster than odd (“marked”) numbers.

In the case of multi-digit numbers, another linguistic property known as the inversion property is of particular importance. German two-digit number words are inverted: The unit digit is articulated first, followed by the decade digit (e.g., 25 is “fünfundzwanzig”—“five-and-twenty”). In other languages, like English or Polish, the structures of the number word systems are comparable to the Arabic number notation, i.e., the decade digit is articulated first and followed by the unit digit. The inversion property in German can lead to problems with transcoding, i.e., children mixing up units and decades when writing numbers on dictation (Zuber et al., 2009). Transcoding in inverted number systems seems to demand more working memory and executive function resources (Imbo et al., 2014). Inversion can also affect symbolic arithmetic in German-speaking children (Göbel et al., 2014). Effects of inversion on arithmetic performance (Van Rinsveld et al., 2015) and magnitude judgments (Van Rinsveld et al., 2016) can also be observed in adults. Comparing the German number word system with the Japanese (i.e., a more transparent) number word system, German-speaking children show not only more transcoding errors in general, but a specific pattern of transcoding errors reflecting the unit-decade inversion property in their number word system (Moeller et al., 2015). Additionally, due to the inversion property, German-speaking

participants automatically pay more attention to the unit digit of a given two-digit number word, as this digit is articulated first, while English speaking participants tend to pay more attention to the decade digit (Nuerk et al., 2005a). This effect is present in different modalities: In non-inverted languages, decades seem to play a greater role in processing than units, regardless of whether numbers are presented visually or auditorily (Macizo and Herrera, 2008, Exp. 3; Macizo and Herrera, 2010). This prioritizing of either the unit or decade digit might influence participants’ performance in number processing tasks in which units play a decisive role. Parity judgement is clearly one of those tasks, because only the unit (parity) is relevant for answering correctly.

However, not only the composition of number words influences number processing, but also the grammatical number (singular, plural) assigned to a number (Roettger and Domahs, 2015). Most languages, like English and German, follow simple rules regarding grammatical number: While 1 is associated with singular, all other numbers are associated with plural. In Polish, grammatical number rules in verbal inflection are more complex: while 1 is associated with singular, 2, 3, and 4 are associated with plural, but 5–9 are again associated with singular. The grammatical number for multi-digit numbers follows analogous rules. All numbers ending in 1 (as well as teens and full decades) are associated with singular, all numbers ending in 2, 3, or 4 are associated with plural and all numbers ending in a number from 5–9 are associated with singular again. For example, 24 is associated with plural (“There are 24.”), and 27 is associated with singular (“There is 27.”). These grammatical number rules cause an incongruence between numerical and grammatical number for numbers associated with singular grammatical number, which could have an impact on their representation and processing. Nevertheless, such influences have not yet been demonstrated and this point needs to be treated as a rather tentative prediction.

Altogether, linguistic factors are expected to influence number processing, and, therefore, to affect response speed for parity judgment. Thus, we expect reaction times for the examined numerical properties to differ cross-linguistically. Due to these linguistic influences, our initial account might not accurately depict the effect of the odd-even continuum for different language groups.

Other Factors Influencing Numerical Judgments: Magnitude and Word Frequency

Numerous studies investigating numerical processing point out that numerical magnitude and frequency of a given number word in natural language affect decision times on numerical stimuli. These effects can be observed both in parity and magnitude judgments. Therefore, we consider them as potentially influencing our results, despite being irrelevant to the postulated parity continuum account.

First of all, processing of numbers is affected by their magnitude. Larger numbers are associated with longer reaction times in number comparison tasks (i.e., the size effect; Moyer and Landauer, 1967). In parity judgement tasks, the size effect

has also been reported (e.g., Gevers et al., 2006), but the evidence is less conclusive (e.g., Dehaene et al., 1993; Verguts et al., 2005). Further, numerical magnitude is also mapped onto space (i.e., *Spatial Numerical Association of Response Codes*, SNARC effect). Namely, in bimanual decision tasks, reactions to small/large magnitude numbers are faster on the left/right hand side (Dehaene et al., 1993; Fias, 2001; Nuerk et al., 2005b, for auditory stimuli). For two-digit numbers, SNARC effects can be found depending on the magnitude of the whole number (Tlauka, 2002), the unit magnitude (Huber et al., 2015) and the decade magnitude (Dehaene et al., 1993). Thus, the magnitude of a whole number as well as the magnitude of the constituents of a multi-digit number have an impact on number processing. In order to control for size effects, unit magnitude and decade magnitude were taken into consideration in the present study.

Besides magnitude, the frequency of a number word (Whaley, 1978), can influence number processing. Numbers occurring more often in the natural language are responded to faster than those which are rarer (see e.g., Van Heuven et al., 2014). Nevertheless, this property is not specific to numbers, but rather reflects well-established effects observed in lexical decision tasks, that decisions on words appearing more frequently in a language are faster. To control word frequency effects, log-transformed (\log_{10}) frequency estimates of number words (Gielen et al., 1991) were taken into consideration.

To sum up, properties such as numerical magnitude and word frequency may play a role for numerical judgments, and thus need to be taken into account, although they are not specifically related to the parity continuum account.

The Present Study

The present study aimed at testing all abovementioned numerical and linguistic factors influencing parity judgments of auditorily presented two-digit numbers within one comprehensive account.

Firstly, according to prototypicality, numbers possessing the properties included in our account (i.e., numbers appearing “more odd”/“more even”) are expected to be associated with shorter reaction times. Alternatively, according to an account based on the markedness strength, as we laid out above, odd numbers are linguistically marked and therefore slower. Linguistically, markedness is a strict category, but psychologically, its effects have been shown to be influenced by individual differences, such as handedness (e.g., Huber et al., 2015). Therefore, psychological markedness may also be a graded psychological principle, similar to parity. Still, because markedness leads to slower response times (compared to unmarked concepts), stronger markedness should lead to even slower response times. Overall, the markedness strength account predicts the opposite pattern from the prototypicality account in the case of odd numbers: That increasing oddness (i.e., stronger markedness) will be associated with longer reaction times. On the other hand, for even numbers, increasing evenness (i.e., stronger unmarkedness) according to both prototypicality and the markedness approach, should be associated with shorter reaction times (H1).

Secondly, we expected overall between-language differences in parity decisions. Namely, German speakers should show

significantly shorter reaction times than the other language groups, since unit-decade inversion leads to the digit relevant for parity judgment (the unit) being pronounced first in German (H2.1). Furthermore, specific features of grammatical number in Polish and English (i.e., grammatical number incongruity in the case of more than half of the numbers in Polish), might possibly lead to slower reaction times in Polish than in English speakers, and also slower than German speakers, both due to inversion property in German and grammatical number incongruencies in Polish (H2.2).

Thirdly, linguistic properties might have specific influences on effects within the parity continuum. Effects related to properties of the decade number should be weaker in German speakers, because they can initiate the response before hearing the decade number. Therefore, they can be less affected by decade magnitude or parity congruity (H3.1). Other specific linguistic differences between the English, Polish, and German language groups are expected to influence the processing of parity (H3.2).

METHODS

Participants

A total of 110 participants (71 female; mean age: 21.8 ± 3.9 years; range: 18–40) took part in the experiment. Out of them, 36 participants were native English speakers (23 female, mean age: 20.2 ± 2.2 years; range: 18–31), 36 were native German speakers (23 female, age: 22.2 ± 3.7 years; range: 18–33) and 38 were native Polish speakers (25 female, mean age: 23.0 ± 4.9 years; range: 18–40). All participants were right-handed and had normal or corrected to normal vision. At the time of testing none of our participants had spent more than 1 year in a foreign linguistic environment. Both parents of all participants were native speakers of the same language. None of the participants suffered from any diagnosed learning, psychiatric or neurological disorder. We obtained approval for testing from the local ethics committees at each site of data collection (York, Tuebingen, and Warsaw). Except for two Polish participants who did not specify their field of study, all participants indicated that they were university students or academic staff at the respective testing sites.

All participants gave their written consent to being tested as a participant in this experiment and were free to withdraw from participation at any point. Participants were compensated with credit points, sweets, or with monetary compensation according to local regulations at testing sites.

Materials

The task was a bimanual computerized parity judgment task on two-digit numbers in different notations/modalities (i.e., participants were to decide whether a given number was even or odd), using the “A” (left hand) and “L” (right hand) keys on a keyboard. Response keys were labeled with colored (blue and purple) stickers. The same laptop model was used at each testing site. The task was programmed and data were collected with Presentation 18.1 software (Neurobehavioral Systems Inc., Albany California, USA).

Stimuli were the numbers from 20 to 99 (10–19 in practice sessions). Stimuli were presented as either Arabic numerals, written number words, or auditorily through the computer's speakers. Presentation modality changed after one block and the order of presentation was randomized to avoid order effects. After the first three blocks with different modalities were presented, another three blocks were presented with response-key assignment reversed.

In this article, we decided to focus on results of the auditory presentation, since linguistic effects like unit-decade inversion are expected to be most salient here. It was shown that SNARC/MARC effects can be notation/modality specific (Nuerk et al., 2004) or not (Nuerk et al., 2005b), thus, for simplicity of presentation, here we only report the modality for which we expected to observe most salient effects. Each number was presented 5 times in each block (400 trials in total). Stimuli were pseudorandomized within sets of 80 numbers. Each block was preceded by a practice session, during which accuracy feedback was given and a reminder of the correct response-key assignment was presented in the bottom line of the screen. The practice session consisted of numbers 10–19 and was repeated if an 80% accuracy threshold was not reached. Additionally, a hint card about the response-to-key assignment was placed on the left side next to the laptop and was visible for the duration of the experiment.

For the auditory presentation, each trial started with a black fixation square (25 × 25 pixels), which was presented for a random duration between 175 and 250 ms (jittered in steps of 25 ms). Subsequently, a blurred mask was presented on the screen and stimuli were presented through the speakers of the computer until a response was given or for a maximum duration of 3,000 ms. The next trial started after an inter-stimulus-interval (ISI) of 200 ms. During this time, a gray mask covered the screen. The volume of speakers was set to the maximum level, and this corresponded to the natural loudness of a person speaking next to the participant. The numbers were recorded by female native speakers of the respective languages speaking at a regular tempo. The average length of number words differed between languages: in the case of English it was 3.22 syllables, in Polish 4.94 syllables, and in German 4.11 syllables. All recordings were shorter than 1000 ms and were not adjusted to length in order to keep them natural-sounding.

Procedure

Participants were tested individually. The order of the blocks was counterbalanced across participants. After responding to demographic questions, participants started with the parity judgment task. Both speed and accuracy were stressed in the instructions.

During a break before the change of response-key assignment and after the last block was presented, participants were asked to do paper-pencil tasks that were not further analyzed (LPS-UT3, Kreuzpointner et al., 2013; a speeded 8-mi arithmetic task, as well as AMAS, Hopko et al., 2003). A debriefing sheet was presented on request at the end of testing.

TABLE 1 | Predictors influence on overall response times in all three languages.

	Mean slope (SD)	$t_{(106)}$	p ($q = 0.031$)	d
ODD NUMBERS				
Decade magnitude	8.32 (7.99)	10.80	<0.001	1.04
Unit magnitude	4.24 (5.10)	8.60	<0.001	0.83
Parity congruity	−11.40 (23.30)	−5.05	<0.001	−0.49
Prime number	24.10 (31.70)	7.87	<0.001	0.76
Square	−1.03 (38.10)	−0.28	0.779	−0.03
Multiplication table	7.57 (37.30)	2.10	0.038	0.20
Divisibility by 5	27.70 (39.00)	7.35	<0.001	0.71
Frequency	7.66 (110.00)	0.73	0.468	0.07
EVEN NUMBERS				
Decade magnitude	8.29 (8.56)	10.00	<0.001	0.97
Unit magnitude	2.99 (8.13)	3.80	<0.001	0.37
Parity congruity	3.05 (25.10)	1.25	0.213	0.12
Square	−11.70 (55.40)	−2.18	0.031	−0.21
Multiplication table	15.40 (39.20)	4.06	<0.001	0.39
Power of 2	6.36 (49.40)	1.33	0.185	0.13
Divisibility by 4	−10.90 (27.10)	−4.17	<0.001	−0.40
Frequency	9.35 (240.00)	0.41	0.682	0.04

q , FDR-corrected alpha level; d , Cohen's d . Significant predictors are marked with bold.

Data Preparation and Analysis

Data Exclusion

Results from practice sessions were not analyzed. The average error rate was 6.34% and errors were not analyzed due to the ceiling effect in a simple task such as parity judgement. Only reaction times associated with correct responses were further analyzed. Due to technical problems, data from three participants (one per language) were not recorded. Reaction times shorter than 200 ms were treated as anticipations and were excluded. Additionally, reaction times that deviated more than ± 3 standard deviations from a participant's mean were excluded sequentially with an update of the mean and standard deviation computation after a trial was excluded until no further exclusions occurred (see e.g., Cipora and Nuerk, 2013 for the same procedure). Due to an error in the programming procedure, results of one stimulus (number 97) could not be analyzed. All these procedures resulted in another 6.46% of data exclusions, so that finally, 87.2% of the data were retained for reaction time analysis. In a second step, full decade numbers and tie numbers were discarded from the analysis as they cannot be easily compared to other two-digit numbers (Dehaene et al., 1990; Nuerk et al., 2011a, 2015), and are frequently excluded from stimuli sets (e.g., Moeller et al., 2009; Chan et al., 2011; Macizo and Herrera, 2011). Full decades are highly frequent and processed very fast (Brysbaert, 1995). For instance, bisection tasks are facilitated by including a decade number as one of three numbers in the bisectionable triplet, as well as by staying in the same decade between the first and third number of the triplet (Nuerk et al., 2002; Korvorst et al., 2007; Wood et al., 2008)³.

³It was demonstrated several times that phenomena observed in numerical cognition, such as for example the SNARC effect are highly dependent on the task set (see e.g., Dehaene et al., 1993, Exp. 3; Fias et al., 1996, Exp. 1). Thus, to avoid

TABLE 2 | Predictors influence on response times separately for each language.

Odd numbers	English				German				Polish			
	Mean slope (SD)	$t_{(34)}$	p ($q = 0.038$)	d	Mean slope (SD)	$t_{(34)}$	p ($q = 0.013$)	d	Mean slope (SD)	$t_{(36)}$	p ($q = 0.044$)	d
Decade magnitude	9.87 (4.50)	13.00	<0.001	1.25	−0.18 (5.68)	−0.19	0.851	−0.02	14.80 (4.60)	19.60	<0.001	1.89
Unit magnitude	1.14 (3.70)	1.82	0.078	0.18	5.68 (5.55)	6.05	<0.001	0.59	5.72 (4.54)	7.66	<0.001	0.74
Parity congruity	−17.00 (18.30)	−5.48	<0.001	−0.53	−5.90 (27.40)	−1.27	0.211	−0.12	−8.97 (17.70)	−3.08	0.004	−0.30
Prime number	27.10 (21.10)	7.58	<0.001	0.73	1.60 (32.70)	0.29	0.775	0.03	41.80 (26.70)	9.52	<0.001	0.92
Square	−13.20 (35.10)	−2.23	0.033	−0.22	3.46 (44.40)	0.46	0.647	0.05	6.25 (32.10)	1.18	0.245	0.11
Multiplication table	1.04 (31.20)	0.20	0.845	0.02	−15.10 (33.60)	−2.66	0.012	−0.26	35.90 (27.80)	7.87	<0.001	0.76
Divisibility by 5	38.70 (32.10)	7.13	<0.001	0.69	−1.12 (34.60)	−0.19	0.849	−0.02	44.10 (33.90)	7.92	<0.001	0.77
Frequency	71.20 (90.00)	4.69	<0.001	0.45	25.30 (95.00)	1.57	0.127	0.15	−66.30 (91.00)	−4.43	<0.001	−0.43
Even numbers	Mean slope (SD)	$t_{(34)}$	p ($q = 0.031$)	d	Mean slope (SD)	$t_{(34)}$	p ($q = 0.031$)	d	Mean slope (SD)	$t_{(36)}$	p ($q = 0.006$)	d
Decade magnitude	9.26 (6.25)	8.76	<0.001	0.85	1.10 (7.41)	0.88	0.387	0.09	14.30 (6.40)	13.60	<0.001	1.32
Unit magnitude	5.13 (5.58)	5.43	<0.001	0.53	5.76 (9.77)	3.49	0.001	0.34	−2.08 (6.30)	−2.01	0.052	−0.19
Parity congruity	3.25 (26.90)	0.72	0.479	0.07	5.78 (28.60)	1.19	0.241	0.12	0.79 (20.50)	0.23	0.817	0.02
Square	−15.20 (51.50)	−1.74	0.091	−0.17	−32.80 (60.90)	−3.19	0.003	−0.31	9.30 (46.50)	1.22	0.231	0.12
Multiplication table	25.10 (33.80)	4.38	<0.001	0.42	23.00 (48.00)	2.83	0.008	0.27	−0.21 (29.80)	−0.04	0.966	−0.01
Power of 2	−5.34 (38.10)	−0.83	0.413	−0.08	27.80 (49.20)	3.34	0.002	0.32	−3.04 (53.40)	−0.35	0.731	−0.03
Divisibility by 4	−29.70 (20.30)	−8.65	<0.001	−0.84	1.15 (27.80)	0.25	0.807	0.02	−4.36 (22.90)	−1.16	0.254	−0.11
Frequency	212.00 (120.00)	10.00	<0.001	0.97	−194.00 (210.00)	−5.50	<0.001	−0.53	4.95 (170.00)	0.18	0.859	0.02

q , FDR-corrected alpha level; d , Cohen's d . Significant predictors are marked with bold.

Multiple Regression Analyses (H1)

Within-participant multiple regressions were calculated separately for odd and even numbers. Predictors not specifically related to the parity continuum account were included in both models. These were: (a) Log-transformed (\log_{10}) frequency of a number word estimated by subjective ratings, ranging from 0 to 500 (Gielen et al., 1991)⁴, (b) unit magnitude, (c) decade magnitude, (d) parity congruity. Multiple regressions for even numbers included predictors: being a *square*, being a part of a *multiplication table*, a *power of 2*, as well as being *divisible by 4*. Multiple regressions for odd numbers included predictors: being a *square*, a *prime number*, being part of a *multiplication table*, as well as being *divisible by 5*.

Binary predictors: *parity congruity*, being a *square*, a *prime number*, part of a *multiplication table*, a *power of 2*, as well as being *divisible by 4* and *by 5* were coded as 1 when the particular feature was present, and 0 when they were not. Individual regression slopes (unstandardized beta coefficients) for each predictor served as dependent measures that were further analyzed. Participants' regression slopes for each factor were tested against 0 with a two-sided *t*-test (Lorch and Myers, 1990). Levels of significance were adjusted for multiple comparisons using False Discovery Rate (FDR) correction (Benjamini and Hochberg, 1995). Positive slopes denote longer reaction times for possessing/increasing a given property; negative slopes denote shorter reaction times for possessing/increasing a given property. Regarding our prototypicality hypothesis for the effects of the odd-even continuum (H1), we expected factors which lead numbers to be processed as "more odd" or "more even" to show more negative slopes, that is, to be associated with shorter reaction times.

In order to check for predictor collinearity, we calculated correlations between predictors (See Supplementary Material A). Although in some cases correlations were moderate, they did not exceed 0.57 in any case; thus, it did not raise the problem of collinearity for multiple regressions⁵. However, to check for possible suppression effects (potentially changing the direction of relationships observed within the multiple regression approach), we calculated bivariate correlations between predictors of interest. Averaged within participant bivariate correlations are presented in Supplementary Material B. Furthermore, we checked whether slopes associated with significant effects had the same directions as averaged bivariate correlations. If that was the case, it is mentioned explicitly in the Results section. Note that the setup we used allows calculating the SNARC effect as well. Nevertheless, it was out of the scope of the present study; thus

it is not presented in the following analysis, but it is reported in Supplementary Material C.

Group Comparisons (H2.1 and H2.2; H3)

To investigate whether language groups differed in reaction times (H2.1 and H2.2) and regression slopes (H3), respectively, we calculated one-way ANOVAs. In addition, Bayesian ANOVAs were conducted. Posterior probabilities in favor of the null hypothesis model given the data $p(H_0|D)$ were calculated, with the null hypothesis denoting no between-group differences and the alternative hypothesis denoting between-group differences. Interpretations of posterior probabilities were based on Raftery (as cited in Masson, 2011). All analyses were conducted with R (version 3.3.0; R Core Team, 2018) and JASP (Version 0.8.2; JASP Team, 2017).

Comparing Odd and Even Numbers

To investigate whether the whole sample showed an odd effect (faster mean reaction times for even than for odd numbers in general), a one-way ANOVA was calculated checking for a group difference between even and odd stimuli.

RESULTS

Multiple Regression Analyses (H1 and H3) Whole-Sample Level

Including all participants, multiple linear regression analysis and subsequent *t*-tests revealed significant effects in both odd and even numbers. In odd numbers, *prime number* and *divisibility by 5* showed significant, positive slopes (i.e., were associated with longer reaction times). For even numbers being a *square* and *divisibility by 4* showed negative slopes (i.e., were associated with shorter reaction times). On the contrary, being part of a *multiplication table* was associated significantly with longer reaction times in even numbers (cf. **Table 1**). Interestingly, the bivariate correlation with being part of a multiplication table had the opposite direction from regression slopes, suggesting the presence of suppression effects.

Regarding the other predictors, *parity congruent* numbers were responded to faster than incongruent ones but only in the case of odd numbers. On the other hand, increasing *decade magnitude* and *unit magnitude* were associated with longer reaction times for both odd and even numbers. Frequency was neither significant for odd nor for even numbers (cf. **Table 1**). Unexpectedly, in the case of even numbers the bivariate correlation between unit magnitude and reaction times was negative, suggesting the presence of suppression effects (cf. Supplementary Material B).

Within-Language Group Analyses

Subsequently, regression slopes were tested against zero separately for each language group. Checking whether given effects were observed within each language group was a necessary prerequisite for comparing language groups as a next step.

English

For odd numbers, *t*-tests on regression slopes revealed significant effects of being a *prime number*, being a *square*, and *divisibility by*

such possible effects, the entire number range was used in the task, and then full decades and tie numbers were excluded *post-hoc*.

⁴We consider this database as the standard in the field of numerical cognition and a better proxy of the real frequency when a cross-lingual design is applied.

⁵Intercorrelations between predictors are considered problematic if they exceed .80. Another value indicating collinearity is the Variance Inflation Factor (VIF), which should not exceed 10 (see e.g., Field et al., 2012; p. 292-293). Some authors recommend even lower acceptable VIF values. To the best of our knowledge, the most conservative threshold is 3. In our case the maximal VIF values were 2.04 and 2.53 for odd and even numbers respectively.

5. Being a prime number and divisibility by 5 were associated with longer reaction times, whereas being a square was significantly associated with shorter reaction times (cf. **Table 2**). In the case of even numbers, being part of a *multiplication table* was associated with longer reaction times, while being a *square* and *divisibility by 4* resulted in shorter reaction times (cf. **Table 2**). Notably, in the case of being part of a multiplication table, the bivariate correlation had an opposite direction suggesting the presence of suppression effects (cf. Supplementary Material B).

As regards the other predictors, *parity congruent* numbers were responded to faster than incongruent ones but only in the case of odd numbers. On the other hand, increasing *decade magnitude* was associated with longer reaction times for both odd and even numbers. Increasing *unit magnitude* was significantly associated with increasing reaction times only for even numbers. Frequency was significant for both odd and even numbers (cf. **Table 2**). More frequent numbers were responded to slower than less frequent ones.

German

For odd numbers, results of *t*-tests on regression slopes revealed a significant association of being part of a *multiplication table* with shorter reaction times (cf. **Table 2**). In the case of even numbers, being part of a *multiplication table* or a *power of 2* were significant positive predictors, meaning possessing these numerical properties was associated with longer reaction times. In addition, being a *square* led to shorter reaction times (cf. **Table 2**).

As regards the other predictors, *parity congruity* and *decade magnitude* were not significant. On the other hand, increasing *unit magnitude* was significantly associated with increasing reaction times for both odd and even numbers. Frequency was significant only in even numbers (cf. **Table 2**). More frequent numbers were responded to faster than less frequent ones.

Polish

For odd numbers, being a *prime number*, being part of a *multiplication table*, and *divisibility by 5* were significant positive predictors, meaning possessing them was associated with longer reaction times. Nevertheless, the bivariate correlation between being part of a multiplication table and reaction time was negative (cf. Supplementary Material B), suggesting possible suppression effects. For even numbers, none of the specific predictors reached significance (cf. **Table 2**).

As regards the other predictors, *parity congruent* numbers were responded to faster but only in the case of odd numbers. Increasing *decade magnitude* was associated with longer reaction times for both odd and even numbers, while increasing *unit magnitude* was associated with longer reaction times only in odd numbers. Increasing frequency was related with shorter reaction times only in odd numbers (cf. **Table 2**).

Between-Group Differences in Mean Reaction Time (H2.1 and H2.2) and the Odd Effect

To address H2.1 and H2.2, and to check for a presence of the odd effect, a mixed design 3 (language) \times 2 (parity) ANOVA

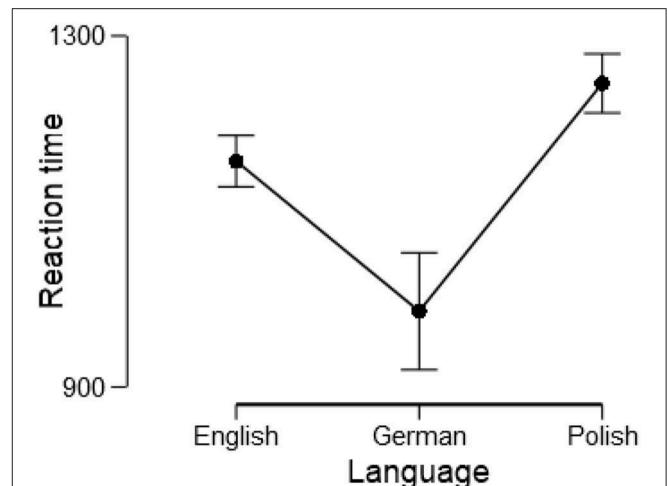


FIGURE 2 | Mean reaction times with 95% confidence interval for the English, German, and Polish language group.

was conducted. There was a robust effect of Language, $F_{(2,214)} = 68.04$, $p < 0.001$, $\eta_p^2 = 0.39$ (cf. **Figure 2**). Post hoc comparison revealed that all groups different significantly from each other ($ps < 0.001$). Interestingly, there was no main effect of number parity, $F_{(1,214)} = 0.24$, $p = 0.628$, $\eta_p^2 < 0.01$ indicating absence of the odd effect⁶. The interaction parity \times language was also not significant, $F_{(2,214)} = 0.02$, $p = 0.979$, $\eta_p^2 < 0.01$, thus the odd effect was not modulated by language.

Between-Group Comparisons (H3)

For odd numbers, ANOVAs testing for group differences in regression slopes revealed significant differences between language groups for *prime number*, *being part of a multiplication table*, and *divisibility by 5* (cf. **Table 3**). For even numbers, ANOVAs revealed significant differences between language-groups for factors *being a square*, *being part of a multiplication table*, *power of 2*, and *divisibility by 4*. To support these results, Bayesian ANOVAs were calculated, as well (cf. **Table 3**). As regards the other predictors, groups did not differ in *parity congruity*. On the other hand, there were differences as regards the effects of *decade magnitude*, *unit magnitude*, and frequency for both odd and even numbers (cf. **Table 3**).

DISCUSSION

Results of a parity judgment task with two-digit numbers in three language groups (English, German, and Polish) were analyzed regarding numerical properties for odd and even numbers in

⁶Hines et al. (1996) found a more pronounced odd effect in males when numbers were presented as dot patterns, whereas in females the effect was more pronounced in case of number words. Therefore, in their study the effect and its direction depended on presentation format. They did not use auditory modality, therefore direct replication within our study was not possible. However, to follow-up on their results we checked for gender effects in our data. We did not find a significant Gender \times Parity interaction, $F < 1.00$; $p = 0.842$. There was also no Gender \times Parity \times Language interaction $F < 0.01$; $p > 0.999$.

TABLE 3 | Predictors influence on response times as compared between three languages.

	$F_{(2,104)}$	p	η_p^2	$p(H0 D)$	BF01	$p(H1 D)$	Interpretation	Post-hoc
ODD NUMBERS								
Decade magnitude	86.41	<0.001	0.62	0.000	0.000	1.000	Very strong for H1	All groups differ
Unit magnitude	11.56	<0.001	0.18	0.001	0.001	0.999	Very strong for H1	E differs from G and P
Parity congruity	2008.00	0.139	0.04	1.000	2276.000	0.000	Very strong for H0	Not applicable
Prime number	21.13	<0.001	0.29	0.000	0.000	1.000	Very strong for H1	All groups differ
Square	2799.00	0.065	0.05	0.999	1201.000	0.001	Very strong for H0	Not applicable
Multiplication table	24.96	<0.001	0.32	0.000	0.000	1.000	Very strong for H1	P differs from G and E
Divisibility by 5	19.46	<0.001	0.27	0.000	0.000	1.000	Very strong for H1	G differs from E and P
Frequency	21.73	<0.001	0.30	0.000	0.000	1.000	Very strong for H1	P differs from G and E
EVEN NUMBERS								
Decade magnitude	35.03	<0.001	0.40	0.000	0.000	1.000	Very strong for H1	All groups differ
Unit magnitude	11.04	<0.001	0.18	0.002	0.002	0.998	Very Strong For H1	P differs from G and E
Parity congruity	0.43	0.652	0.01	1.000	8113.000	0.000	Very strong for H0	not applicable
Square	6506.00	0.002	0.11	0.059	0.063	0.941	Positive for H1	G differs from E and P
Multiplication table	5363.00	0.006	0.09	0.133	0.154	0.867	Positive for H1	P differs from G and E
Power of 2	5321.00	0.006	0.09	0.139	0.162	0.861	Positive for H1	G differs from E and P
Divisibility by 4	16.60	<0.001	0.24	0.000	0.000	1.000	Very strong for H1	E differs from G and P
Frequency	50.33	<0.001	0.49	0.000	0.000	1.000	Very strong for H1	All groups differ

H0, null hypothesis or no between-group difference, H1, alternative hypothesis or group difference, E, English, P, Polish, G, German. Significant predictors are marked with bold.

order to verify the parity continuum account and language differences in parity processing. We observed robust language differences in overall reaction times thus confirming hypotheses H2.1 and H2.2. Hypotheses regarding direction of mean slopes (H1), as well as linguistic differences regarding mean slopes (H3) could partially be confirmed and were partially contradicted, which will be discussed below. It was not straightforward to test the tentative account directly, because the postulated categories are neither fully independent of each other nor fully nested (e.g., odd squares are neither a subset of numbers divisible by 5, nor is it the other way around). Instead, after controlling for the effects of *parity congruity*, *unit*, and *decade magnitude*, as well as *frequency*, we compared the regression slopes for numerical properties potentially influencing the perceived parity with the parity continuum account. Those numerical properties comprised being a *prime number*, a *square*, part of a *multiplication table*, and being *divisible by 5* for odd numbers, as well as being a *square*, part of a *multiplication table*, a *power of 2*, and being *divisible by 4* for even numbers.

Conclusions for the Tentative Account

The fundamental assumption that time needed for parity judgments differs considerably depending on numerical properties was confirmed by the data. However, the strict order postulated by neither by the prototypicality nor the markedness strength account was not fully captured.

For odd numbers, being a *prime number* and being *divisible by 5* was related to systematically longer reaction times. Despite having a robust impact on reaction times, the pattern of results was not in line with the prototypicality account's predictions that increasing easiness of division would make numbers subjectively less odd and thus associated with longer reaction

times. Accordingly, primes would be responded to fastest, and numbers divisible by 5, slowest. The results were also not in line with predictions driven from the markedness strength account that “most odd” numbers, i.e., the primes would be responded to slowest.

This surprising result suggests that different factors might play a role in parity decisions and thus the account considering one dimension only (i.e., easiness of division) seems too simple to explain all numerical influences. Being *part of a multiplication table* and *being a square* were not significant predictors of reaction times (cf. **Figure 3**) in the whole sample analyses.

In the case of even numbers, being *part of a multiplication table*, *divisibility by 4*, and *being a square* significantly predicted reaction times. Expectedly, divisibility by 4 and being a square were associated with shorter reaction times. This can be due to the easiness of division dimension we introduced. However, *being part of a multiplication table* was associated with longer reaction times. This surprising result needs further investigation in future studies, as numbers that are part of a multiplication table are used more often than those which are not. On the other side, being part of a multiplication table does not determine a number's parity status, and possibly, accessibility of respective division facts might be harmful for parity processing, so that one needs to verify whether division facts are specifically related to divisibility by 2. Notably, the direction of the slope was different than the direction of bivariate correlation, thus suggesting the presence of suppression effects in the case of this predictor. This should also be addressed in future studies. The effect of *being a power of 2* was not significant. Nevertheless, slopes related to being a power of 2 were estimated based on two numbers only (32 and 64), so it may be that if one would use more repetitions of these numbers in a more specific setup, it would

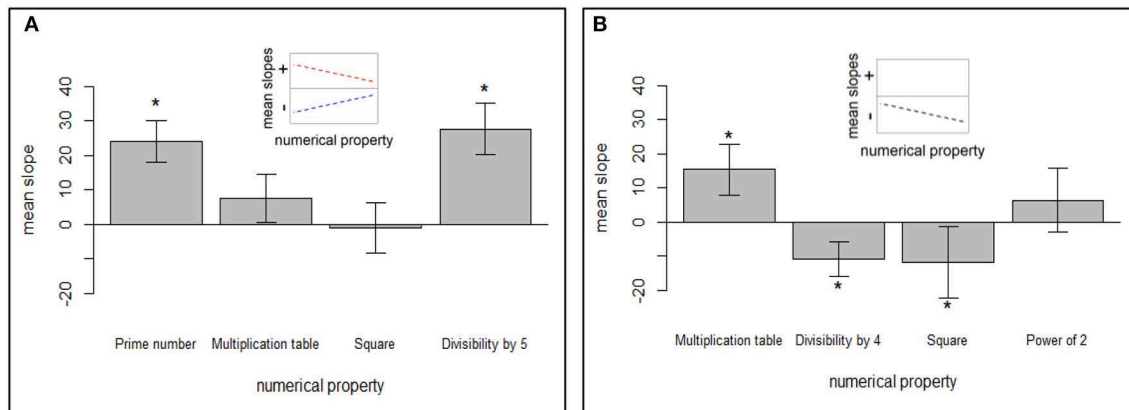


FIGURE 3 | Mean slopes with 95% confidence intervals for numerical properties of (A) odd and (B) even numbers across groups; *indicating significance after correcting for multiple comparisons. Small panels represent predictions regarding the overall tendency we expected to observe. For odd numbers, according to the prediction derived from the prototypicality account, the bars in this figure should be arranged in an increasing order (schematically represented by blue line in the small panel). In case of prediction driven from the markedness strength account, the tendency is the opposite—bars should represent a decreasing order (as schematically depicted by red line in the small panel). For even numbers, there was only one prediction driven by the prototypicality account: decreasing order of bars (as schematically depicted in the small panel).

be possible to observe a more consistent effect. Despite the suboptimal design for investigating the effect of being a power of 2, we decided to retain this predictor in our model, because we had strong predictions regarding these numbers, and we thought that excluding it could potentially decrease the overall model fit.

Linguistic Effects as Limitations and Refinements for the Tentative Account (H2.1 and H2.2; H3)

Our hypotheses regarding differences in mean overall reaction time between language groups were confirmed: German speaking participants reacted the fastest, while Polish speaking participants the slowest (H2.1 and H2.2). In the case of German participants, reaction times were shortest mostly due to the inversion property—the decisive unit number was heard first so that participants could start to give the response, or at least prepare it. This effect was indeed observed and reaction times were the fastest in German participants, despite the considerably larger syllable length of number words in German than in English. On the other hand, Polish speakers were the slowest, which might be either due to the fact that Polish number words were longest, or due to specific grammatical number properties. Note that the point in time at which specific number words are recognized differs across languages. For instance, to accurately categorize the number 91 in Polish, the decisive syllable “je,” being the first syllable of the number of units, appears in the fifth position of the number word “dziewiećdziesiąt jeden,” whereas in German the decisive “ein” syllable appears in the first position of the number word “einundneunzig.”

Furthermore, due to the inversion property, one might also expect that numerical properties will affect German speakers to a lesser extent than English and Polish speakers. Interestingly, this

was true only in the case of odd numbers. In the case of even numbers, German speakers were highly affected by numerical properties, but Polish speakers were not (cf. **Figure 4**).

The overall effects of being a *prime number* and *divisibility by 5* were driven only by English and Polish speakers but were not present in German speakers. Recognizing whether a given number is a prime requires processing a whole two-digit number. Thus, the absence of an effect in German can be explained by the fact that German speakers make their parity decisions based on units only and can simply ignore the following decade number. However, the lack of an effect of *divisibility by 5* in German is puzzling. Divisibility by 5 can be accessed based on the number of units; thus, its effect should be present in German speakers as well.

Interestingly, for odd numbers being *part of a multiplication table* was a significant predictor in German and Polish speakers. Nevertheless, the direction of the effect was opposite (shorter reaction times in German and slower in Polish), and the effects canceled each other out. This means that the prototype hypothesis was corroborated in Polish. Being part of a multiplication table makes a number less typically odd (than for instance being a prime number) and therefore, RTs are slower. In contrast, for German speakers, the markedness hypothesis seems to be true in that these “less odd” numbers are faster, because they are less marked. We did not hypothesize this result. Two explanations are possible. First, maybe markedness is particularly pronounced in German, possibly because marked adjectives are often obvious, because negating prefixes are particularly common. The second hypothesis refers to multiplication learning. Possibly, learning multiplication tables is not so highly overlearned anymore (our personal anecdotal impression from many studies is that many elementary school teachers do not like the drill associated with it) and therefore the effect of prototypicality is less pronounced than

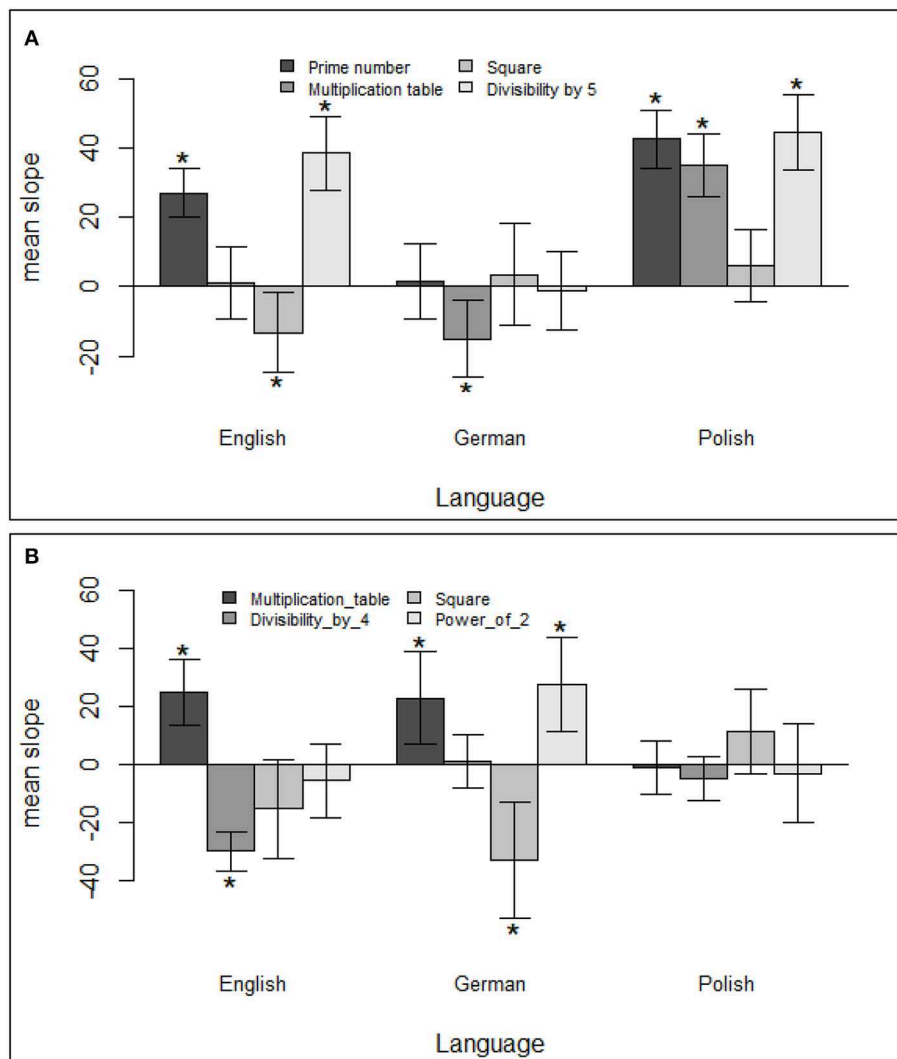


FIGURE 4 | Mean slopes with 95% confidence intervals for numerical properties of (A) odd and (B) even numbers in the English, German, and Polish group; *indicating significance after correcting for multiple comparisons.

in Poland. This needs to be tested in future cross-cultural studies in which the ease of multiplication table activation is also assessed in the same participants. Another important difference as regards the prototypicality account refers to the inversion effect in German. Because the unit is spoken first, the whole multiplication number does not need to be processed before the parity decision is initiated (when one hears “seven-and-twenty,” he or she can initiate the response when he or she hears “seven”). Therefore, the activation of the identity of the whole number may be less or later. Consequently, the influence of prototypicality as derived from multiplication attributes of the whole number may be weaker in German.

English, in which no effects were found, may be a mix between Poland and Germany as regards markedness and prototypicality effects. However, we wish to note that the direction of the effect in Polish might be due to suppression. Finally, the effect of *being*

square was significant only in English speakers. Since this refers to 4 numbers only (9, 25, 49, 81), we would not wish to make any strong claims at this first study on the subject.

In the case of even numbers, none of the numerical predictors reached significance in Polish speakers. In the case of English and German speakers, the effect of being *part of a multiplication table* was significant and went in the same direction (but suggests suppression effects in German). On the other hand, it seems that the overall effect of *divisibility by 4* was driven by English speakers only, while the overall effect of being a *square* was driven by German speakers only. The effects in English and German can be explained by both markedness and prototypicality as outlined above. The null effects in Polish come as a surprise but could be due to a weaker role of markedness in the Polish language that could already partially explain the effects for odd numbers. Again, this explanation is tentative and requires further specialization.

Overall, while some language effects pointed in the hypothesized direction, others were pointing in the opposite direction. Possible causes are linguistic, educational, and cultural differences, different salencies of the prototype and markedness strength hypotheses in different languages, but also methodological issues like a small number of stimuli in some categories and possible collinearities.

To begin with, in the introduction, we outlined the prototype and the markedness strength hypotheses. For even numbers, these hypotheses predicted the same things. Multiplication attributes should lead to faster RT. For odd numbers, they predicted opposite patterns. While the prototype account predicted faster RTs for more prototypical odd numbers (e.g., prime numbers), the markedness strength account predicted longer RTs for such numbers, because they are psychologically more marked and therefore processed even slower.

The predictions for even numbers (*divisibility by 4*, *being a square number*) followed the prototype and markedness hypotheses. Only being part of a multiplication table was not in the expected direction. It is conceivable that this effect is due to complex suppression effects, because divisibility by 4 and being a square number overlap with multiplication effects. This tentative explanation seems to be supported by observation that bivariate correlations went in the opposite direction than multiple regression slopes.

The predictions for odd numbers are more complicated than we had anticipated. Some of the results seem to favor the prototype hypothesis, while other seem to favor the markedness strength hypothesis. Our presumption is that both hypotheses may be valid and that their saliency depends on linguistic, educational and cultural properties. For instance, being a *prime number* prolonged RT in English and Polish, thus favoring a markedness strength account for this attribute. However, it did not prolong RT in German, probably because the parity decision in German could be finished before the whole number (and hence the identity of the prime number) was finished. Similarly, the effect of *being part of a multiplication table* went in opposite directions in German and Polish. While the faster RTs in German seemed to favor a markedness strength account for this attribute, the slower RTs in Polish seemed to favor the prototypicality account. However, markedness saliency induced by parity is similar in both languages, because odd is a negation of even in both languages (“*ungerade*” vs. “*gerade*” in German, “*nieparzysty*” vs. “*parzysty*” in Polish). Therefore, there might be other linguistic or cultural or educational factors, which may favor the markedness strength account in German and the prototypicality account in Polish, which we do not yet fully understand. All in all, while some patterns observed with regard to the odd numbers like the different effects of prime numbers can be explained based on the available accounts, other differences, such as multiplication table influences cannot be easily explained. We wish to acknowledge, however that because of collinearities and nested effects (prime numbers are by definition not part of the multiplication tables), suppression effects and therefore a methodological explanation rather than a theoretical remains possible.

Effects of Congruity, Size, and Frequency

The factor *parity congruity* was included to investigate unit-decade congruity effects in even and odd numbers. For odd numbers, participants responded slower to incongruent stimuli at the whole-sample level, as well as in the Polish and English group, but not in the German group. This fits with the inversion property of German language, because it is easier for German speakers to ignore the task-irrelevant decade number being presented as second. For English and Polish the interfering decade number is spoken first before the response-relevant unit digit, while for German the response-relevant unit digit is spoken first and the answer can in principle be initiated before the decade digit is even presented. Interestingly, for even numbers, parity congruity did not influence reaction times, either on the whole-sample level, or in any of the three individual language groups. An explanation for this unexpected effect is tentative. However, we need to keep in mind that responding to even numbers is faster (*odd effect*, Hines, 1990). Evenness is the unmarked pole of the parity representation and is as such the more dominant ground form, which is easier to access and more salient. It is conceivable that there is an equivalent to the global precedence in global-local research (Navon, 1977; but see Kimchi, 1992) in that there is a precedence for processing even numbers, which receive less interference by odd numbers than vice versa (at least for auditory numbers and with a balanced stimuli set such as we used).

Decade and *unit magnitude* affected reaction times significantly for both odd and even numbers at the whole-sample level. Increasing magnitude was associated with longer reaction times. This size effect (Moyer and Landauer, 1967)—the bigger the number, the slower the response—differed significantly between language groups in both odd and even numbers.

The decade magnitude effect was present in both odd and even numbers at the whole sample level as well as in English and Polish, but not in German speakers. Again, this might be due to the inversion property of German.

The results regarding the unit magnitude are also fairly straightforward. It was apparent for both odd and even numbers at the whole sample level. Interestingly, it was present in German speakers for both odd and even numbers, which shows that magnitude effects are present in this language group but are further modulated by linguistic properties for both unit and decade digits in the expected direction. Nevertheless, the effect of unit magnitude was not present for odd numbers in English speakers or even numbers in Polish speakers. Again, the processing of unit magnitude begins later in English and Polish (because there is no inversion) and it might be weaker for the less salient odd numbers than for the more salient even numbers. In sum, the findings for decade and unit magnitude effects for different languages and for different parities largely mimic those observed for the parity congruity effect. Generally, the influence of the unit is larger in German (because of inversion), while the influence of the decade is larger in English and Polish. If there are further differences between parities, magnitude is more likely activated for even parities than for odd parities.

The *frequency* of number words was controlled for by including it as a factor in the analysis. For both odd and even numbers, *frequency* was not significant on a whole-sample level. Nevertheless, the effect of frequency was robust in both odd and even numbers in English; however, surprisingly larger frequency was associated with longer reaction times.

In the case of odd numbers in Polish and even numbers in German the effect was in line with predictions, so that higher frequency was associated with shorter reaction times. The effect was not present for odd numbers in German or even numbers in Polish. At the current stage, we do not have an explanation for this interaction between language and parity with regard to frequency effects.

GENERAL CONCLUSIONS

To begin with the hypotheses regarding, linguistic differences were robustly reflected in our results. First of all, German speakers were less affected by decade magnitude than English and Polish speakers. However, the effect of decade magnitude was not totally eliminated in this group. Namely, this group revealed some effects which depended on decade magnitude, such as responding faster to odd numbers that were part of a multiplication table. Such effects can be only explained by the decade number being at least partly processed, because such information can be extracted only when overall numerical magnitude is processed. On the other hand, inconsistent grammatical number did not play a robust role in parity decisions in Polish speakers. This might be due to the fact that numerical processing was not framed in any linguistic context in the present experiment—participants were presented with numbers only, not embedded in any additional phrasing.

Effects of multiplicativity and other numerical variables on parity could be observed but were not always consistent. For even numbers, being a *square* and *divisibility by 4* led to shorter reaction times, i.e., made a number “more even.” To give an example: 64 (square and divisible by 4) is “more even” than 62 (not a square and not divisible by 4). Note, being part of a *multiplication table* was associated significantly with longer reaction times in even numbers in the regression analysis (cf. **Table 1**). However, the bivariate correlation with being part of a multiplication table had the opposite direction from regression slopes, suggesting the presence of suppression effects. So at least in the raw correlations, 42 (part of the multiplication table: 6×7) would be more even than 46. However, this relation is more tentative than for being a *square* and *divisibility by 4*, because of the reversal of the slope in the multiple regression.

For odd numbers, the interpretation is more difficult, because the prototype and markedness account predict opposing response patterns and our cross-lingual analysis suggest that both may play a role. In line with the outlined markedness strength account, for odd numbers we observed a gradual decrease in response time, starting from prime numbers to numbers that

are part of a multiplication table and finally squares. So, 23 (being a prime number) was slower than 27 (being part of the multiplication table (3×9), which was slower than a square number (25, but see below). In contrast to those multiplicativity attributes, divisibility by 5 rather followed the prototypicality, as it slowed down responses: (e.g., 45 was slower than 47 or 49, when all other factors (prime, square number) were partialled out)—this is in line with the idea that numbers divisible by 5 are not typical odd numbers and are therefore slower to be categorized as odd. In sum, for odd numbers, we can say that multiplication attributes influence parity decisions strongly and significantly. However, it seems that we are looking at two opposing effects here, markedness strength and prototypicality, which compete with each other. Therefore, a simple order according to RT like for even numbers cannot be provided so easily.

All in all, however, the current data suggest that not all numbers are equally odd or equally even. Several aspects of two-digit numbers, their multiplicativity, their parity congruity, and in some languages their frequency influence parity categorization. Dependent on language, culture, education and predictor, sometimes less prototypical numbers of a category are slower responded to, corroborating the prototypicality account, while in other cases more marked numbers (and in the case of odd numbers, therefore more prototypical numbers) are slower responded to. Which account is most salient for which language and which attribute is an endeavor for future research. However, we wish to acknowledge that methodological constraints like collinearities or having few members of a category might also have influenced the results and produced suppression and interaction effects. This is not a fault of the current study, as we used all two-digit numbers above 19, but instead an inherent attribute of our numerical system. For instance, there are just two even square numbers between 20 and 99, namely 36 and 64 (note that both of them are divisible by four and one of them is also a power of 2). Of course, 2 members in one category is much less than anybody would have liked. Therefore, independent replications of our results are necessary to see how stable the results for a given language will be.⁷

Nevertheless, although not every single multiplicativity predictor (especially for small stimulus groups and high collinearity) may prevail in a replication, the present results quite clearly show that the parity judgments are not all the same. There are some consistent findings that unit and decade magnitude, parity congruity, but also some attributes like being a prime number or being divisible by 4 influence parity decisions in a fairly consistent way across languages. Therefore, we believe it is fair after this study to conclude that not all even/odd numbers are psychologically equally even or odd, respectively. However,

⁷Note that pairwise matching is probably impossible, because this study suggests that so many different attributes (decade magnitude, unit magnitude, parity congruity, frequency and different multiplicativity attributes) may influence reaction times. These would need to be controlled for pairwise matching, which is probably impossible.

we also have to acknowledge that the mechanisms responsible for making numbers more even or odd in a given language or culture need to be better studied and understood in the future.

ETHICS STATEMENT

The study was approved by the ethic committee of the Medical Faculty of the University of Tuebingen. It got further approval at other data collection sites (University of York, Department of Psychology and University of Warsaw, Department of Psychology).

AUTHOR CONTRIBUTIONS

KC, MS, KL, SG, FD, MH, and H-CN designed the study. LH, M-LS collected the data. LH, KC, M-LS, and MS analyzed the data. LH, KC, and MS wrote the manuscript. All authors read and commented on and corrected the manuscript.

REFERENCES

- Armstrong, S. L., Gleitman, L. R., and Gleitman, H. (1983). What some concepts might not be. *Cognition* 13, 263–308. doi: 10.1016/0010-0277(83)90012-4
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.* 57, 289–300.
- Brysbaert, M. (1995). Arabic number reading: on the nature of the numerical scale and the origin of phonological recoding. *J. Exp. Psychol.* 124, 434–452. doi: 10.1037/0096-3445.124.4.434
- Chan, W. W. L., Au, T. K., and Tang, J. (2011). Exploring the developmental changes in automatic two-digit number processing. *J. Exp. Child Psychol.* 109, 263–274. doi: 10.1016/j.jecp.2011.01.010
- Cipora, K., and Nuerk, H.-C. (2013). Is the SNARC effect related to the level of mathematics? No systematic relationship observed despite more power, more repetitions, and more direct assessment of arithmetic skill. *Q. J. Exp. Psychol.* 66, 1974–1991. doi: 10.1080/17470218.2013.772215
- Dehaene, S., and Akhavein, R. (1995). Attention, automaticity, and levels of representation in number processing. *J. Exp. Psychol.* 21, 314–326. doi: 10.1037/0278-7393.21.2.314
- Dehaene, S., Bossini, S., and Giraux, P. (1993). The mental representation of parity and number magnitude. *J. Exp. Psychol.* 122, 371–396. doi: 10.1037/0096-3445.122.3.371
- Dehaene, S., Dupoux, E., and Mehler, J. (1990). Is numerical comparison digital? Analogical and symbolic effects in two-digit number comparison. *J. Exp. Psychol.* 16, 626–641. doi: 10.1037/0096-1523.16.3.626
- Domahs, F., Moeller, K., Huber, S., Willmes, K., and Nuerk, H.-C. (2010). Embodied numerosity: implicit hand-based representations influence symbolic number processing across cultures. *Cognition* 116, 251–266. doi: 10.1016/j.cognition.2010.05.007
- Dowker, A., and Nuerk, H.-C. (2016). Linguistic influences on mathematics. *Front. Psychol.* 7:1035. doi: 10.3389/fpsyg.2016.01035
- Fias, W. (2001). Two routes for the processing of verbal numbers: evidence from the SNARC effect. *Psychol. Res.* 65, 250–259. doi: 10.1007/s004260100065
- Fias, W., Brysbaert, M., Geypens, F., and D'Ydewalle, G. (1996). The importance of magnitude information in numerical processing: evidence from the SNARC effect. *Math. Cogn.* 2, 95–110. doi: 10.1080/135467996387552
- Field, A., Miles, J., and Field, Z. (2012). *Discovering Statistics Using R*. London: Sage publications.
- French, D. (2005). Double, double, double. *Math. School* 34, 8–9.
- Gevers, W., Ratinckx, E., De Baene, W., and Fias, W. (2006). Further evidence that the SNARC effect is processed along a dual-route architecture:

ACKNOWLEDGMENTS

We would like to thank all participants. This research was funded by a grant from the DFG (NU 265/3-1) to H-CN supporting KC and MS, and from the National Science Center (NCN), Poland (2014/15/G/HS6/04604) to MH supporting KL. KC, MS, and H-CN are further supported by the LEAD Graduate School & Research Network (GSC1028), which is funded within the framework of the Excellence Initiative of the German federal and state governments. We acknowledge support by Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of University of Tuebingen. Finally, we thank our assistants who helped with data collection and language proofreading the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.01081/full#supplementary-material>

- evidence from the lateralized readiness potential. *Exp. Psychol.* 53, 58–68. doi: 10.1027/1618-3169.53.1.58
- Gielen, I., Brysbaert, M., and Dhondt, A. (1991). The syllable-length effect in number processing is task-dependent. *Attent. Percept. Psychophys.* 50, 449–458. doi: 10.3758/BF03205061
- Göbel, S. M., Moeller, K., Pixner, S., Kaufmann, L., and Nuerk, H.-C. (2014). Language affects symbolic arithmetic in children: the case of number word inversion. *J. Exp. Child Psychol.* 119, 17–25. doi: 10.1016/j.jecp.2013.10.001
- Guttman, L. (1968). A general nonmetric technique for finding the smallest coordinate space for a configuration of points. *Psychometrika* 33, 469–506. doi: 10.1007/BF02290164
- Hines, T. M. (1990). An odd effect: lengthened reaction times for judgments about odd digits. *Mem. Cognit.* 18, 40–46. doi: 10.3758/BF03202644
- Hines, T. M., Herman-Jeglinska, A., Bednarek, D., and Grabowska, A. (1996). Sex differences in the processing of odd and even numbers. *Acta Neurobiol. Exp.* 56, 263–266.
- Hopko, D. R., Mahadevan, R., Bare, R. L., and Hunt, M. K. (2003). The abbreviated math anxiety scale (AMAS) construction, validity, and reliability. *Assessment* 10, 178–182. doi: 10.1177/1073191103010002008
- Huber, S., Klein, E., Graf, M., Nuerk, H.-C., Moeller, K., and Willmes, K. (2015). Embodied markedness of parity? Examining handedness effects on parity judgments. *Psychol. Res.* 79, 963–977. doi: 10.1007/s00426-014-0626-9
- Imbo, I., Vanden Bulcke, C., De Brauwier, J., and Fias, W. (2014). Sixty-four or four-and-sixty? The influence of language and working memory on children's number transcoding. *Front. Psychol.* 5:313. doi: 10.3389/fpsyg.2014.00313
- Kimchi, R. (1992). Primacy of wholistic processing and global/local paradigm: a critical review. *Psychol. Bull.* 112, 24–38. doi: 10.1037/0033-2909.112.1.24
- Klein, E., Moeller, K., Willmes, K., Nuerk, H.-C., and Domahs, F. (2011). The influence of implicit hand-based representations on mental arithmetic. *Front. Psychol.* 2:197. doi: 10.3389/fpsyg.2011.00197
- Korvorst, M., Nuerk, H.-C., and Willmes, K. (2007). The hands have it: number representations in adult deaf signers. *J. Deaf Stud. Deaf Educ.* 12, 362–372. doi: 10.1093/deafed/enm002
- Kreuzpointner, L., Lukesch, H. and Horn, W. (2013). *Leistungsprüfsystem 2. LPS-2*. Göttingen: Hogrefe.
- Lingoes, J. C., and Roskam, E. E. (1973). A mathematical and empirical analysis of two multidimensional scaling algorithms. *Psychometrika* 38:93.
- Lorch, R. F., and Myers, J. L. (1990). Regression analyses of repeated measures data in cognitive research. *J. Exp. Psychol.* 16:149. doi: 10.1037/0278-7393.16.1.149
- Macizo, P., and Herrera, A. (2008). The effect of number codes in the comparison task of two-digit numbers. *Psicológica* 29, 1–34.

- Macizo, P., and Herrera, A. (2010). Two-digit number comparison: decade-unit and unit-decade produce the same compatibility effect with number words. *Can. J. Exp. Psychol.* 64, 17–24. doi: 10.1037/a0015803
- Macizo, P., and Herrera, A. (2011). Cognitive control in number processing: evidence from the unit-decade compatibility effect. *Acta Psychol.* 136, 112–118. doi: 10.1016/j.actpsy.2010.10.008
- Masson, M. E. (2011). A tutorial on a practical Bayesian alternative to null-hypothesis significance testing. *Behav. Res. Methods* 43, 679–690. doi: 10.3758/s13428-010-0049-5
- Moeller, K., Fischer, M. H., Nuerk, H.-C., and Willmes, K. (2009). Sequential or parallel decomposed processing of two-digit numbers? Evidence from eye-tracking. *Q. J. Exp. Psychol.* 62, 323–334. doi: 10.1080/17470210801946740
- Moeller, K., Zuber, J., Olsen, N., Nuerk, H.-C., and Willmes, K. (2015). Intransparent German number words complicate transcoding – a translingual comparison with Japanese. *Front. Psychol.* 6:740. doi: 10.3389/fpsyg.2015.00740
- Moyer, R. S., and Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature* 215, 1519–1520. doi: 10.1038/2151519a0
- Navon, D. (1977). Forest before trees: the precedence of global features in visual perception. *Cogn. Psychol.* 9, 353–383. doi: 10.1016/0010-0285(77)90012-3
- Nuerk, H.-C., Geppert, B. E., van Herten, M., and Willmes, K. (2002). On the impact of different number representations in the number bisection task. *Cortex* 38, 691–715. doi: 10.1016/S0010-9452(08)70038-8
- Nuerk, H.-C., Iversen, W., and Willmes, K. (2004). Notational modulation of the SNARC and the MARC (linguistic markedness of response codes) effect. *Q. J. Exp. Psychol.* A 57, 835–863. doi: 10.1080/02724980343000512
- Nuerk, H.-C., Moeller, K., Klein, E., Willmes, K., and Fischer, M. H. (2011a). Extending the mental number line. *Zeitschrift für Psychol.* 219, 3–22. doi: 10.1027/2151-2604/a000041
- Nuerk, H.-C., Moeller, K., and Willmes, K. (2015). “Multi-digit number processing - overview, conceptual clarifications, and language influences,” in *Oxford Handbook of Mathematical Cognition*, eds R. Cohen Kadosh and A. Dowker (Oxford: Oxford University Press), 106–139.
- Nuerk, H.-C., Weger, U., and Willmes, K. (2005a). Language effects in magnitude comparison: small, but not irrelevant. *Brain Lang.* 92, 262–277. doi: 10.1016/j.bandl.2004.06.107
- Nuerk, H.-C., Willmes, K., and Fischer, M. H. (2011b). Multi-digit number processing. *Zeitschrift für Psychol.* 219, 1–2. doi: 10.1027/2151-2604/a000040
- Nuerk, H.-C., Wood, G., and Willmes, K. (2005b). The universal SNARC effect: the association between number magnitude and space is amodal. *Exp. Psychol.* 52, 187–194. doi: 10.1027/1618-3169.52.3.187
- Osherson, D. N., and Smith, E. E. (1981). On the adequacy of prototype theory as a theory of concepts. *Cognition* 9, 35–58. doi: 10.1016/0010-0277(81)90013-5
- Posner, M. I., and Keele, S. W. (1968). On the genesis of abstract ideas. *J. Exp. Psychol.* 77, 353–363. doi: 10.1037/h0025953
- R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Roettger, T. B., and Domahs, F. (2015). Grammatical number elicits SNARC and MARC effects as a function of task demands. *Q. J. Exp. Psychol.* 68, 1231–1248. doi: 10.1080/17470218.2014.979843
- Rosch, E. (1975). Cognitive representations of semantic categories. *J. Exp. Psychol.* 104, 192–233. doi: 10.1037/0096-3445.104.3.192
- Rosch, E., and Lloyd, B. B. (Eds.). (1978). *Cognition and Categorization*, Vol. 1. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rosch, E., Simpson, C., and Miller, R. S. (1976). Structural bases of typicality effects. *J. Exp. Psychol.* 2, 491–502. doi: 10.1037/0096-1523.2.4.491
- Sasanguie, D., Defever, E., Van den Bussche, E., and Reynvoet, B. (2011). The reliability of and the relation between non-symbolic numerical distance effects in comparison, same-different judgments and priming. *Acta Psychol.* 136, 73–80. doi: 10.1016/j.actpsy.2010.10.004
- Schroeder, P. A., Nuerk, H.-C., and Plewnia, C. (2017). Switching between multiple codes of SNARC-like associations: two conceptual replication attempts with anodal tDCS in sham-controlled cross-over design. *Front. Neurosci.* 11:654. doi: 10.3389/fnins.2017.00654
- Sherman, M. A. (1976). Adjectival negation and the comprehension of multiply negated sentences. *J. Verbal Learn. Verbal Behav.* 15, 143–157. doi: 10.1016/0022-5371(76)90015-3
- Tan, S., and Dixon, P. (2011). Repetition and the SNARC effect with one- and two-digit numbers. *Can. J. Exp. Psychol.* 65, 84–97. doi: 10.1037/a0022368
- Tlauka, M. (2002). The processing of numbers in choice-reaction tasks. *Aust. J. Psychol.* 54, 94–98. doi: 10.1080/00049530210001706553
- Van Heuven, W. J., Mandera, P., Keuleers, E., and Brysbaert, M. (2014). SUBTLEX-UK: a new and improved word frequency database for British English. *Q. J. Exp. Psychol.* 67, 1176–1190. doi: 10.1080/17470218.2013.850521
- Van Rinsveld, A., Brunner, M., Landerl, K., Schiltz, C., and Ugen, S. (2015). The relation between language and arithmetic in bilinguals: insights from different stages of language acquisition. *Front. Psychol.* 6:265. doi: 10.3389/fpsyg.2015.00265
- Van Rinsveld, A., Schiltz, C., Landerl, K., Brunner, M., and Ugen, S. (2016). Speaking two languages with different number naming systems: what implications for magnitude judgments in bilinguals at different stages of language acquisition? *Cogn. Process.* 17, 225–241. doi: 10.1007/s10339-016-0762-9
- Verguts, T., Fias, W., and Stevens, M. (2005). A model of exact small-number representation. *Psychon. Bull. Rev.* 12, 66–80. doi: 10.3758/BF03196349
- Whaley, C. P. (1978). Word–nonword classification time. *J. Verbal Learn. Verbal Behav.* 17, 143–154. doi: 10.1016/S0022-5371(78)90110-X
- Wood, G., Willmes, K., Nuerk, H.-C., and Fischer, M. H. (2008). On the cognitive link between space and number: a meta-analysis of the SNARC effect. *Psychol. Sci.* 50, 489–525.
- Zuber, J., Pixner, S., Moeller, K., and Nuerk, H.-C. (2009). On the language specificity of basic number processing: transcoding in a language with inversion and its relation to working memory capacity. *J. Exp. Child Psychol.* 102, 60–77. doi: 10.1016/j.jecp.2008.04.003

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Heubner, Cipora, Soltanlou, Schlenker, Lipowska, Göbel, Domahs, Haman and Nuerk. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Corrigendum: A Mental Odd-Even Continuum Account: Some Numbers May Be “More Odd” Than Others and Some Numbers May Be “More Even” Than Others

OPEN ACCESS

Approved by:

Frontiers Editorial Office,
Frontiers Media SA, Switzerland

*Correspondence:

Krzysztof Cipora
krzysztof.cipora@uni-tuebingen.de

†These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 18 July 2019

Accepted: 29 July 2019

Published: 16 August 2019

Citation:

Heubner L, Cipora K, Soltanlou M,
Schlenker M-L, Lipowska K,
Göbel SM, Domahs F, Haman M and
Nuerk H-C (2019) Corrigendum: A
Mental Odd-Even Continuum
Account: Some Numbers May Be
“More Odd” Than Others and Some
Numbers May Be “More Even” Than
Others. *Front. Psychol.* 10:1860.
doi: 10.3389/fpsyg.2019.01860

Lia Heubner^{††}, Krzysztof Cipora^{1,2*†}, Mojtaba Soltanlou^{1,2}, Marie-Lene Schlenker¹,
Katarzyna Lipowska³, Silke M. Göbel⁴, Frank Domahs⁵, Maciej Haman³ and
Hans-Christoph Nuerk^{1,2,6}

¹ Department of Psychology, University of Tuebingen, Tuebingen, Germany, ² LEAD Graduate School and Research Network, University of Tuebingen, Tuebingen, Germany, ³ Department of Psychology, University of Warsaw, Warsaw, Poland,

⁴ Department of Psychology, University of York, York, United Kingdom, ⁵ Institute for German Linguistics, University of Marburg, Marburg, Germany, ⁶ Leibniz-Institut für Wissensmedien, Tuebingen, Germany

Keywords: parity judgment, markedness, numerical properties, prototypicality, cross-linguistic comparisons

A Corrigendum on

A Mental Odd-Even Continuum Account: Some Numbers May Be “More Odd” Than Others and Some Numbers May Be “More Even” Than Others

by Heubner, L., Cipora, K., Soltanlou, M., Schlenker, M.-L., Lipowska, K., Göbel, S. M., et al. (2018). *Front. Psychol.* 9:1081. doi: 10.3389/fpsyg.2018.01081

There is an error in the Acknowledgments statement. In the original article, we acknowledge the funder “NCN”. The correct name for the Funder is the “National Science Center (NCN), Poland.”

In addition, the correct funding number for the National Science Center (NCN), Poland is “2014/15/G/HS6/04604.”

The authors apologize for these errors and state that they do not change the scientific conclusions of the article in any way. The original article has been updated.

Copyright © 2019 Heubner, Cipora, Soltanlou, Schlenker, Lipowska, Göbel, Domahs, Haman and Nuerk. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Developing Mental Number Line: Does Its Directionality Relate to 5- to 7-Year-Old Children's Mathematical Abilities?

Lauren S. Aulet* and Stella F. Lourenco*

Department of Psychology, Emory University, Atlanta, GA, United States

OPEN ACCESS

Edited by:

Krzysztof Cipora,
Universität Tübingen, Germany

Reviewed by:

Koen Luwel,
KU Leuven, Belgium
Carrie Georges,
University of Luxembourg,
Luxembourg

*Correspondence:

Lauren S. Aulet
lauren.s.aulet@emory.edu
Stella F. Lourenco
stella.lourenco@emory.edu

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 29 January 2018

Accepted: 14 June 2018

Published: 06 July 2018

Citation:

Aulet LS and Lourenco SF (2018)
The Developing Mental Number Line:
Does Its Directionality Relate to 5-
to 7-Year-Old Children's Mathematical
Abilities? *Front. Psychol.* 9:1142.
doi: 10.3389/fpsyg.2018.01142

Spatial representations of number, such as a left-to-right oriented mental number line, are well documented in Western culture. Yet, the functional significance of such a representation remains unclear. To test the prominent hypothesis that a mental number line may support mathematical development, we examined the relation between spatial-numerical associations (SNAs) and math proficiency in 5- to 7-year-old children. We found evidence of SNAs with two tasks: a non-symbolic magnitude comparison task, and a symbolic "Where was the number?" (WTN) task. Further, we found a significant correlation between these two tasks, demonstrating convergent validity of the directional mental number line across numerical format. Although there were no significant correlations between children's SNAs on the WTN task and math ability, children's SNAs on the magnitude comparison task were negatively correlated with their performance on a measure of cross-modal arithmetic, suggesting that children with a stronger left-to-right oriented mental number line were less competent at cross-modal arithmetic, an effect that held when controlling for age and a set of general cognitive abilities. Despite some evidence for a negative relation between SNAs and math ability in adulthood, we argue that the effect here may reflect task demands specific to the magnitude comparison task, not necessarily an impediment of the mental number line to math performance. We conclude with a discussion of the different properties that characterize a mental number line and how these different properties may relate to mathematical ability.

Keywords: spatial-numerical associations, mental number line, SNARC, mathematical ability, development

INTRODUCTION

Interest in the spatial nature of numerical representations can be traced back as early as 1880 to Francis Galton's work on "number forms." In this work, Galton demonstrated that individuals visualized numbers in spatial format, albeit in an idiosyncratic manner across individuals (Galton, 1880). In subsequent, now seminal work, Dehaene et al. (1993) documented systematic associations, the so-called SNARC (spatial-numerical association of response codes) effect, among

Western participants who made parity (odd/even) judgments to Arabic numerals using left and right response keys. In this study, participants responded faster to smaller numbers when using the left key and responded faster to larger numbers when using the right key, providing evidence of a left-to-right spatial representation of number (see also Zorzi et al., 2002; Fischer et al., 2003; Wood et al., 2008), often referred to as the mental number line¹. In the 25 years since the publication of this work, there have been ongoing efforts to understand the ontogenetic and phylogenetic origins of such a number line (for alternative perspectives, Gevers et al., 2003; Abrahamse et al., 2016). Because this initial work dealt primarily with symbolic numerical stimuli (Dehaene et al., 1993), and because it has since been shown that there is cross-cultural variation in the direction of these effects (Zebian, 2005; Shaki et al., 2009), it was hypothesized that the mental number line arose from experience with linguistic conventions (i.e., reading and writing). Although cultural experience certainly modulates the directionality of one's mental number line (Shaki et al., 2009; McCrink et al., 2014), recent research using non-symbolic stimuli demonstrates directional effects in non-human animals (Drucker and Brannon, 2014; Rugani et al., 2015) and preliterate children (Patro and Haman, 2012; de Hevia et al., 2014), though the specific orientation of these directional effects may vary (Cooperrider et al., 2017; Gazes et al., 2017).

Yet, there are open questions regarding both the developmental trajectory and functional significance of a mental number line. To address such questions, it is important to acknowledge that the mental number line is a heterogeneous phenomenon. There are both directional and non-directional properties of the mental number line (for review, see Cipora et al., 2015). The directional property of the mental number line refers specifically to the orientation of the mapping between numbers and space. For example, Westerners orient numbers left-to-right, in contrast to non-Westerners who may orient numbers right-to-left (Shaki et al., 2009). Among the non-directional properties is the type of spatial scaling (see recent meta-analysis, Schneider et al., 2018). For example, there is a large literature on the extent to which the mental number line can be considered linear as opposed to compressive (e.g., logarithmic; Siegler and Opfer, 2003; Dehaene et al., 2008; Lourenco and Longo, 2009). This literature has provided evidence for a developmental shift from compressive to linear number representations (Siegler et al., 2009; Opfer et al., 2016) as well as for a relation between spatial scale and math development. In particular, children's performance on a variety of math tasks has been found to be positively associated with the linearity of their representations (Siegler and Booth, 2004;

Booth and Siegler, 2006; Sasanguie et al., 2013). However, compared to the non-directional properties of the mental number line, far less work has concerned the development and function of the directional nature of the mental number line, particularly the potential relation between directionality and math competence. Thus, the present study focused specifically on the development and function of SNAs as a way to ask whether directionality, like spatial scaling, affords any benefit to math development.

Assessing SNAs in children has proven difficult, as most tasks designed for adults utilize relatively advanced numerical judgments (i.e., parity), bimanual responses, and/or reaction time (RT) data, which may be unsuitable for children or pose interpretative challenges when testing developmental populations. To address these challenges, Opfer et al. (2010) implemented a search task in which they found that 4-year-olds were more accurate at locating an item in a series of horizontally arranged containers when the containers were numbered from left-to-right, as opposed to right-to-left (see also Opfer and Furlong, 2011). Similarly, they found that children counted a series of horizontally arranged items from left-to-right more often than from right-to-left. Though this work suggests that young children are sensitive to culturally specific counting practices (Shaki et al., 2012), it is less clear whether their performance was driven by access to a left-to-right mental number line.

Other researchers have devised tasks that minimize the demands imposed on children while also maintaining similarity to those used with adults in an effort to address questions about the developmental continuity of SNAs. For example, van Galen and Reitsma (2008) adopted magnitude, instead of parity, judgments (i.e., Is the presented Arabic numeral smaller or larger than 5?) and found evidence of SNAs in 7- to 9-year-old children, as has been shown in adults. In this study, children responded faster to smaller numbers when using the left key and responded faster to larger numbers when using the right key, consistent with a left-to-right mental number line. Patro and Haman (2012) further addressed the methodological challenges of testing children with a non-symbolic magnitude comparison task (i.e., a comparison of dot displays) and unimanual responses. In their study, 2- to 4-year-olds judged which of two simultaneously presented arrays contained more items by selecting the array with the larger numerosity. They found that children were significantly faster when the target array was on the right side of space than on the left. They also showed a similar, albeit weaker, effect when children selected the array that was smaller in numerosity, such that children were somewhat quicker when the target array was on the left. Taken together, these findings provide support for a directional mental number line that emerges early in development, when experience with reading and mathematics is minimal. However, even with some evidence for directional number representations beginning in infancy (Bulf et al., 2016), the reliability of these effects remains unknown. Moreover, and crucially, little is known about the potential functional significance of such number representations. In particular, we can ask whether the directionality of the mental number line has

¹In the present study, we define the mental number line as a representation of number in which numerical value is associated with a spatial location. The spatial properties of the mental number line include both directionality (i.e., the specific directional orientation of the line) and scaling (i.e., the spatial intervals between numerical values). We define spatial-numerical associations (SNAs) as the behavioral manifestations of these properties, though the current work concerned directionality in particular. To this end, we describe SNAs here as behavioral manifestations of the directionality of the mental number line.

any impact on one's understanding of mathematical concepts and reasoning.

Does the Directionality of the Mental Number Line Have Functional Significance?

One prominent hypothesis is that a mental number line functions to support mathematical development and understanding (Opfer et al., 2010; Fischer and Shaki, 2014). On this view, stronger SNAs would be accompanied by better math ability. Although evidence of SNAs in non-human animals would seem to argue against this perspective (since non-human animals never develop formal math skills), it remains possible that, at least in humans, mathematical reasoning recruits numerical representations shared by humans and non-human animals. This possibility is buttressed by research on the approximate number system (ANS; Feigenson et al., 2004), which has found that the ANS, a non-verbal system of number representation that humans, similarly, share with non-human animals (Cantlon and Brannon, 2006; Rugani et al., 2008; Agrillo et al., 2011) and that is operational early in human development (Xu and Spelke, 2000; Cordes and Brannon, 2008), predicts formal math abilities when assessed concurrently (Libertus et al., 2011; Bonny and Lourenco, 2013) and longitudinally (Halberda et al., 2008; Starr et al., 2013).

Yet, existing research on the relation between SNAs and mathematical ability conducted on adult participants does not provide strong evidence for a directional mental number line that benefits mathematical reasoning. In particular, one line of evidence suggests a negative relation, such that participants with math difficulties showed stronger SNAs than participants without math difficulties (Hoffmann et al., 2014a). Likewise, non-mathematicians (i.e., doctoral students in the humanities and social sciences) have been found to show stronger SNAs than mathematicians (i.e., doctoral students in mathematics; Cipora et al., 2016). Other studies, however, have reported no relation between SNAs and math competence in adulthood (Bull et al., 2013; Cipora and Nuerk, 2013), demonstrating inconsistency in the extant research with adults.

The challenge with using adult participants to test whether SNAs may be related to mathematical reasoning is that adults have a mature system of mathematics at their disposal. Although there is certainly variability in the math exposure adults experience, most adults in Western society have learned a variety of mathematical concepts and are capable of performing computations on numbers and variables. For these reasons, one might ask whether the spatial instantiation of mathematical concepts would be of greater utility earlier in development, when these concepts are initially acquired and remain difficult to grasp. Recent evidence suggests that early spatial skills such as mental rotation predict later math development (Lauer and Lourenco, 2016; Verdine et al., 2017), suggesting a potentially broader role for spatial representations in math learning. As highlighted above, however, SNAs have been difficult to assess in children, impeding the study of their potential utility at the critical early stages of math development. To date, only a few studies have assessed the relation between SNAs and math

ability in childhood, and, as with adults, the findings have been inconsistent.

Hoffman et al. (2013) tested the relation between SNAs and math ability in 5-year-old children who completed color and magnitude judgments on Arabic numerals (similar to the task used by van Galen and Reitsma, 2008). Children also completed measures of numerical competence (e.g., verbal counting and digit writing). Although children showed evidence of an SNA when judging the color of Arabic numerals, extending the finding of van Galen and Reitsma (2008) to younger children, individual performance did not correlate with any measure of numerical competence. By contrast, when SNAs were indexed using the magnitude judgment task, there was a positive correlation between children's SNAs and some measures of numerical competence. Importantly, however, the SNA effect, at the group-level, was not significant, raising questions about the reliability of children's SNAs on the magnitude judgment task in this study and, thus, the reported links with numerical competence.

In other work, Bachot et al. (2005) examined a group of 16 children from ages 7 to 12 years with visuospatial deficits, some of whom also had dyscalculia, a mathematical learning disability. They found that children with visuospatial and mathematical deficits did not exhibit SNAs on a symbolic magnitude judgment task, whereas a control group of children, matched for age and gender, did. However, because children with visuospatial and math deficits were not differentiated in the analyses, it is unclear whether the lack of a significant SNA was driven by poor math ability or visuospatial deficits. More recently, Gibson and Maurer (2016) used a similar symbolic magnitude comparison task to assess SNAs in 6- to 8-year-old typically-developing children. They found that 6-year-olds, like the children in the study of Hoffman et al. (2013), did not show an SNA on this task. However, in older children (7- and 8-year-olds), the SNA effect was significant, but this effect did not relate to math performance, as assessed by a standardized measure of symbolic math ability (for similar results, see Schneider et al., 2009). In a study of 8- to 11-year-olds ($n = 55$), Georges et al. (2017b) assessed children's SNAs for symbolic numerals with a parity judgment task. They observed a positive relation between children's SNAs and performance on the arithmetic, but not visuospatial, subtest of a standardized, speeded math exam, such that children with stronger SNAs exhibited better arithmetic skill. However, this relation was only observed in younger children (8–10 years old), in contrast with Gibson and Maurer (2016), who observed no relation between SNAs and performance on a standardized assessment of symbolic math in children of similar age.

Taken together, the extant data from studies with children suggest that a mental number line, with consistent directionality, may be present as young as preschool age (for work with infants, see Bulf et al., 2016). However, the evidence in support of a relation between an early emerging directional mental number line and mathematical development is mixed, with several open questions following from the existing findings. For example, it remains unknown how associations between symbolic and non-symbolic representations of number affect the link between SNAs and math competence, given that a directional mental number line with symbolic numerals may not be present until 7 years

of age. As in studies with adults, it is also unclear whether the relation between SNAs and math competence in children depends on the type of math ability assessed such as whether arithmetic computations are performed exactly or approximately.

Present Study

In the present study, we assessed children's performance on two SNA tasks, as well as on multiple measures of early mathematical competence. Children were between 5 and 7 years of age, an age range in which formal math instruction has only recently begun and, thus, when even basic mathematical concepts and operations may not yet be mastered. The two measures of SNAs assessed the directionality of children's number representations (mental number line) using different judgments and either symbolic or non-symbolic stimuli. One SNA task was a non-symbolic version of the magnitude comparison task (Patro and Haman, 2012), and the other SNA task was a novel "Where was the number?" (WTN) task with Arabic numerals (adapted from Aulet et al., 2017). In this task, children simply viewed a number on a screen, memorized its location, and, after a short delay, placed the number back in its original location. As noted previously, existing research using Arabic numerals in a magnitude comparison task has not provided evidence of SNAs until approximately 7 years of age. However, the WTN task required no explicit judgment of magnitude, but rather, only memory for the location of a number that had appeared in a random location on screen, which we reasoned might allow for earlier detection of the directionality of the mental number line with symbolic stimuli. Moreover, the differences in stimuli and task requirements across these two tasks provided a strong test of construct validity. In other words, if a stable, directional mental number line underlies performance on both tasks, then children's performance on the two tasks should be correlated.

We also examined the relation between children's performance on the SNA tasks and their mathematical ability (see **Table 1** for all tasks used in the present study). Because math is not a monolithic concept and, crucially, because the link between SNAs and math ability may depend on the type of math that is assessed, we included multiple measures of early math ability. One possibility is that the understanding of the abstract nature of number would benefit from a grounding in space (for review, see Lourenco et al., 2018). Indeed, the directionality of the mental number line could support the understanding of number as an abstract concept with ordinal structure (Cipora et al., 2015). Another possibility is that this directionality could provide support for the enactment of arithmetic operations, as suggested by the spatial-directional biases associated with addition and subtraction, known as "operational momentum" (McCrink et al., 2007; Pinhas and Fischer, 2008; see also: Klein et al., 2014; Holmes et al., 2016). In particular, it has been suggested that addition and subtraction elicit rightward and leftward movement, respectively, along the mental number line. The mental number line could provide a concrete method for instantiating the arithmetic operations by distinguishing addition and subtraction in terms of directional movement and perhaps by supporting implementation of the computation (Booth and Siegler, 2008; Siegler, 2016). In the present study,

we tested children on tasks designed to target these areas of emerging math competency. In particular, children completed a task that required coordination of numerical information across modalities (vision and audition), as an assessment of children's abstract number representations. Specifically, this task served as a test of abstractness since children must "abstract" across the perceptual information to achieve a common number representation across format. Children also completed two measures of symbolic arithmetic (one approximate and one exact) that assessed competence with arithmetic computation on numerals.

Furthermore, we assessed the internal consistency of all SNA and math tasks in the present study to ensure that any observed relations (or lack thereof) between SNAs and math ability could not be attributed to poor task reliability. Although assessment of reliability is especially critical when utilizing an individual differences approach, previous studies on the relation between SNAs and math ability have infrequently reported the reliability of measures. Finally, to assess the specificity of the link between children's SNAs and their developing math abilities, we included several tasks to control for general cognitive functioning. These tasks assessed verbal proficiency (WJ–Picture Vocabulary subtest; Woodcock et al., 2001a), verbal working memory (WJ–Auditory Working Memory subtest; Woodcock et al., 2001b), and spatial short-term memory (Spatial Memory subtest of the Kaufman–Assessment Battery for Children (K-ABC); Kaufman and Kaufman, 1983). Previous studies reporting a relation between SNAs and math ability have not controlled for general cognitive functioning, leaving open the possibility that other abilities shared by a directional mental number line and math tasks could account for the reported relation. Here we directly addressed this possibility.

MATERIALS AND METHODS

Participants

Sixty-six children (28 female) between the ages of 5 and 7 years of age ($M = 74.65$ months, $SD = 9.62$ months) from the greater Atlanta area participated in this study. One child was excluded from the analyses for failing to complete multiple tasks. Caregivers provided written informed consent on behalf of their children. All children received stickers throughout the session to maintain motivation, as well as a small gift at the end of the session for participating in the study. Experimental procedures were approved by the local ethics committee.

Tasks and Procedure

"Where Was The Number?" Task

In the "WTN" task (adapted from Aulet et al., 2017), children viewed an Arabic numeral (1–9) presented in black font within a rectangle [white fill with black outline; 915×495 pixels ($24.29 \text{ cm} \times 13.10 \text{ cm}$)]. At the start of each trial, a number appeared at a random location within the rectangle (the "whiteboard"). Children were instructed to press a virtual button located at the bottom of the screen ("START") once they felt they had sufficiently memorized the location of the

TABLE 1 | Description of the tasks used in the present study and the constructs they were designed to assess.

Task	Construct	Number format	Calculation type
'Where was the number?' (WTN)	Spatial-Numerical Associations	Symbolic	N/A
Magnitude Comparison	Spatial-Numerical Associations	Non-symbolic	N/A
Approximate Cross-modal Arithmetic (ACA)	Math Ability	Non-symbolic	Approximate
Approximate Symbolic Arithmetic (ASA)	Math Ability	Symbolic	Approximate
WJ – Calculation*	Math Ability	Symbolic	Exact
WJ – Auditory Working Memory*	Verbal Working Memory	N/A	N/A
WJ – Picture Vocabulary*	Verbal Proficiency	N/A	N/A
K-ABC – Spatial Memory*	Spatial Short-Term Memory	N/A	N/A

For numerical tasks, we further specify the number format used and the calculation type assessed. *Standardized tasks.

number. When the start button was pressed, the number disappeared and an image of a dry-erase marker appeared, presented centrally. This was done to ensure that children did not visually fixate on the original location of the number and that all children initiated their responses from the same starting location. Next, children tapped the image of the dry-erase marker, which then disappeared. Children then made their responses by tapping the location on the whiteboard where they believed the number previously appeared. The number appeared at the tapped location. Adjustments to responses could be implemented by tapping and dragging the number to a new location. When satisfied with the placement of the number, children pressed a virtual button located at the bottom of the screen ("Done!") to confirm their response and they proceeded immediately to the next trial.

Presentation of numbers and duration of response window were untimed. Children completed 72 trials in total (each number presented eight times each). To ensure that children remained attentive throughout the task, trials were split into four blocks, each consisting of 18 trials (each number presented twice; random order).

Magnitude Comparison Task

Following Patro and Haman (2012), children completed "more" and "less" conditions of our magnitude comparison task (order counterbalanced across children). In the more condition, children were asked to judge which of two dot arrays was larger in numerosity. In the less condition, children were asked to judge which of two arrays was smaller in numerosity. Following Gebuis and Reynvoet (2011), non-numerical properties in these arrays, such as element size and convex hull, were varied across trials to ensure no systematic relation between these properties and numerosity. Numerical arrays (13.72 cm × 13.72 cm) were arranged horizontally on screen, each below an image of a Star Wars character (BB-8 and R2D2).

In each condition, children completed three practice trials in which they were given corrective feedback. In the practice trials, the two arrays differed in numerosity by a 1:2 ratio (i.e., arrays of 4 vs. 8, 5 vs. 10, and 8 vs. 16). In the test trials, the two arrays differed in numerosity by a 4:5 ratio (i.e., arrays of 4 vs. 5, 8 vs. 10, 12 vs. 15, and 16 vs. 20). Children completed 16 test trials (each ratio presented four times) in each condition, for a total of 32 test trials. On half of the trials, arrays were presented

in the congruent position, with the numerically smaller array presented on the left and the numerically larger array presented on the right. On the other half of the trials, arrays were presented in the incongruent position, with the numerically smaller array presented on the right and the numerically larger array presented on the left. Following previous research (Mazzocco et al., 2011), arrays were visible for 1,200 ms before being occluded. Arrays remained occluded until children responded and then proceeded to the next trial (1,500 ms ISI). All responses were made on a touchscreen.

Approximate Cross-Modal Arithmetic (ACA) Task

In the ACA task (adapted from Barth et al., 2008), we measured the extent to which children's representations of number were modality independent by testing their ability to perform addition and subtraction across displays of dots and sequences of tones. At the beginning of each condition, children completed a familiarization phase as well as two practice trials. In the familiarization phase, children were shown an example animation in which the appearance (addition condition) or disappearance (subtraction condition) of blue dots, one-by-one, was paired with a tone. After this animation, a new array of blue dots was displayed (dots presented simultaneously) and was then occluded by a matching blue occluder. They were told that if they listened carefully, they would hear more blue dots "appear" or "disappear," at which time they heard a sequence of tones. The experimenter then asked the child whether there would be more or less dots behind the occluder than before. If children answered correctly in these demonstrations ("more" for addition and "less" for subtraction), then the experimenter proceeded to the practice trials. If children answered incorrectly, then the experimenter repeated the previous animations.

Children were given two practice trials in which an array of blue dots (19.30 cm × 13.72 cm) was displayed on the left side of the screen (dots presented simultaneously) and was then occluded. After occlusion, children heard a sequence of tones, representing the appearance/disappearance of blue dots. While the blue occluder remained on screen, an array of red dots appeared (19.30 cm × 13.72 cm) on the right side of the screen that was then covered by a matching red occluder (arrays and occluders were matched for luminance). Children were asked whether there were more dots behind the blue or red occluder. After their response, the experimenter removed the occluders

to reveal both arrays, providing children corrective feedback. In both practice trials, blue and red dots differed by a 1:2 ratio (one trial with more blue dots and one trial with more red dots). Following a response, the experimenter advanced to the next trial.

After the practice trials, children completed 12 test trials (randomly ordered). In these trials, blue and red dots differed by one of three ratios: 4:5, 4:6, or 4:7. Children completed four trials of each ratio (two trials in which the blue array was more numerous and two trials in which the red array was more numerous). As in Barth et al. (2008), element size was held constant on all trials of this task. Importantly, though, reliance on non-numerical cues was not likely to account for performance on this task, as success on the task required addition/subtraction of elements across vision and audition (in which the cues differed). The same tone (duration = 15 ms) was used in all tone sequences. This tone was repeated multiple times in each sequence, presented in an irregular rhythm. In the addition condition, final set sizes (dots plus tones), ranged from 16 to 54 ($M = 35$). In the subtraction condition, final set sizes (dots minus tones) ranged from 7 to 30 ($M = 16$). The duration of the sequence of tones ranged from 1.70 to 3.70 s. Although these durations were likely too fast to allow for consistent counting, children were told at the start of the task not to count the individual items and any child who displayed evidence of counting was immediately instructed not to do so. This procedure was used to ensure that all children added or subtracted the sequences using the same approximation strategy.

At the start of each test trial, an array of blue dots (against a solid black background) was displayed on the left side of the screen for 3 s. Then, the blue array was occluded and remained occluded for 6 s while the sequence of tones played. Following this presentation, an array of red dots was displayed on the right side of the screen for 3 s and was then occluded. In all trials of the ACA task, the first display was presented on the left side of the screen and the second display was presented on the right side of the screen. Items were always added to, or subtracted from, the first display². Children were only permitted to respond which array was more numerous once both arrays were occluded. Responses were made using the touchscreen. Immediately after children made their response, the experimenter pressed a key to proceed to the next trial. After children completed the addition condition, the same procedure was completed for the subtraction condition. All children completed the addition condition prior to the subtraction condition. We fixed the trial order in this way because previous research has found that subtraction can be more difficult than addition (Barth et al., 2008) and it has been shown that difficult trials negatively impact performance on subsequent trials (Odic et al., 2014). Thus, to avoid negative carryover effects, addition trials were administered prior to subtraction trials.

²Despite the left-right presentation of displays of the ACA and ASA tasks, an assessment of SNAs on these tasks is not straightforward and, thus, was not implemented. The reason is that the presentation of displays may have elicited additional biases, such as choosing the display that was added to on the addition trials (left display) and choosing the display that was not subtracted from on the subtraction trials (right display). Indeed, some children exclusively chose the left or right display in these tasks (see Results).

Approximate Symbolic Arithmetic (ASA) Task

We assessed children's ability to engage in ASA by requiring them to solve addition and subtraction problems without engaging in exact computation (adapted from Gilmore et al., 2007). In this task, problems were presented verbally along with visual displays containing Arabic numerals. An example problem was: "Sarah has 20 candies in her bag, and then she gets 25 more. John has 30 candies. Which one of them has more candies?" On these problems, the visual displays consisted of cartoon children with accompanying Arabic numerals. Like the ACA task, the first display (e.g., character) was presented on the left side of the screen and the second display was presented on the right side of the screen. Items were always added to, or subtracted from, the first display². Following the reading of the quantities, the corresponding visual displays containing the Arabic numerals were occluded to discourage exact calculation. After the experimenter finished presenting the problem, children responded by pointing to, or naming, the character who they judged as having more candies. Children completed two conditions: addition and subtraction. Within each condition, children completed 12 trials. In the addition condition, final set sizes ranged from 12 to 58 ($M = 30$). In the subtraction condition, final set sizes ranged from 10 to 56 ($M = 28$). Within each trial, final set sizes differed by one of three ratios: 4:5, 4:6, or 4:7. Trials were randomly ordered and untimed. As with the ACA task, the order was fixed to prevent negative carryover effects (Odic et al., 2014), such that the addition condition was administered prior to the subtraction condition.

Exact Symbolic Arithmetic Task

Children completed the Calculation subtest of the Woodcock Johnson (WJ) Tests of Achievement (Woodcock et al., 2001a), a standardized assessment of exact symbolic arithmetic ability. Specifically, the WJ-Calculation test measures participants' ability to perform exact computation using addition, subtraction, multiplication, and division with whole numerals. This test is untimed and administered in paper-and-pencil format following a standard protocol, such that testing is discontinued once six consecutive questions are answered incorrectly.

Control Tasks

Children completed two subtests from the WJ Tests of Achievement and WJ Tests of Cognitive Abilities (Woodcock et al., 2001a,b) that served as controls for general cognitive functioning: verbal proficiency (WJ-Picture Vocabulary) and verbal working memory (WJ-Auditory Working Memory). Children also completed the Spatial Memory subtest from the K-ABC (Kaufman and Kaufman, 1983) as an assessment of spatial short-term memory and to serve as another non-math control task. All control tasks were untimed (for procedural details corresponding to each task, see Kaufman and Kaufman, 1983; McGrew et al., 2007). All control tasks have acceptable reliability, as determined by a split-half procedure: WJ Picture Vocabulary, $r = 0.81$; WJ Auditory Working Memory, $r = 0.96$; K-ABC Spatial Memory, $r = 0.80$ (Kaufman and Kaufman, 1983; McGrew et al., 2007).

General Procedure

All computerized tasks were presented on a Hewlett Packard Compaq Elite 8300 23" all-in-one desktop computer (resolution: 1920 × 1080 pixels). Children were tested individually by an experimenter. Children sat approximately 40 cm from the screen for all computerized tasks. For ease of administration, a fixed order was used such that tasks requiring similar materials were administered consecutively, with computerized tasks preceding paper-and-pencil tasks. Of the computerized tasks, all children first completed the magnitude comparison task, followed by the ACA and ASA tasks (counterbalanced order). Of the paper-and-pencil tasks, all children first completed the WJ–Calculation test. Then, children completed the three control tasks in a randomized order. Given concerns about children's attentiveness across trials, children completed four separate blocks of the WTN task. These blocks were administered at fixed points throughout the session: at the start of the session, as the first task (block 1); after the magnitude comparison task (block 2); after the ACA and ASA tasks (block 3); and after the completion of all standardized tasks, as the last task (block 4). Following the completion of each task, children received a sticker.

RESULTS

Preliminary Analyses

Preliminary analyses showed that scores on all tasks, except for the WTN task, were normally distributed, with skewness statistics within an acceptable range (± 0.60 ; Tabachnick and Fidell, 2001). Scores on the WTN task were transformed (square root transformed) for the correlation analyses reported in the following sections; skewness on the WTN task was in an acceptable range following the transformation.

All tasks yielded acceptable reliabilities ($r_s > 0.52$, Spearman–Brown corrected; see **Table 2**). Reliabilities were calculated for each task using a sample-with-replacement bootstrap technique following Anobile et al. (2016). For each child, the dependent variable (i.e., congruency score for the Magnitude Comparison task, slope for the WTN task, and accuracies for the ACA and ASA tasks) was calculated twice from a random sample of the data (half of the total number of trials for the respective task). This sampling procedure was trial blocked such that equal numbers of each trial type were included (e.g., an equal number of trials for each operation and ratio in the ACA and ASA tasks). We then computed the correlation between the two values, across subjects. This process was repeated 1,000 times and we calculated reliability as the mean correlation for each task.

Children's Performance on the SNA Tasks

WTN Task

Of the total sample, five children were excluded from analyses of the WTN task for failing to complete all four blocks (see **Table 2** for descriptive data). In all analyses of this task, data from the four blocks were combined. Two children were excluded from these analyses due to poor accuracy ($> 2.5 SDs \pm M$), where

accuracy was calculated as the absolute distance between the original location of the number and the child's final placement of the number. The remaining children ($n = 58$) had a mean accuracy of 63.02 pixels [16.67 mm; $SD = 24.42$ pixels (6.43 mm)].

To test for SNAs on this task, the variable of interest was children's bias along the horizontal axis³. For each trial, we calculated the difference between the x-coordinate of children's final placement and the x-coordinate of the number's original location, such that a negative value represented a leftward placement in comparison to the original location, and a positive value represented a rightward placement. For each participant, we then calculated the mean bias for each number and calculated a slope by regressing these values onto their corresponding numerical value. Thus, in this task, a positive slope represents the canonical left-to-right mental number line, as a positive slope denotes a shift from leftward to rightward placement, relative to the number's original position. In other words, just as slopes in the classic SNARC task reflect the extent to which numerical magnitude explains the difference in RTs between left and right hands (Dehaene et al., 1993), slopes in the WTN task reflect the extent to which numerical magnitude explains deviation in the placement of numbers in comparison to the original location.

Consistent with a left-to-right oriented mental number line that applies to symbolic number, children's slopes were significantly greater than zero, $t(57) = 2.02$, $p = 0.048$, $d = 0.265$ (see **Figure 1**), and the majority of children (64%) exhibited a positive slope (binomial test, $p < 0.05$). Children's slopes were not significantly correlated with age, $r(56) = -0.190$, $p = 0.152$, suggesting no relation between symbolic SNAs and age, in a sample of 5- to 7-year-old children. Moreover, children's slopes were not significantly correlated with overall accuracy, $r(56) = 0.134$, $p = 0.316$, suggesting that children's SNAs on this task did not vary as a function of their ability to remember the original location of the number.

Magnitude Comparison Task

Of the total sample, all children were included in the analyses of the more condition; one child was excluded from the analyses of the less condition due to experimenter error. Children's performance (proportion correct) was above the chance level of 0.50 in both conditions (more condition: $M = 0.708$, $SD = 0.147$, $t[64] = 11.41$, $p < 0.001$, $d = 1.42$; less condition: $M = 0.704$, $SD = 0.119$, $t[63] = 13.72$, $p < 0.001$, $d = 1.71$), with no significant difference between the two conditions, $t(63) = 0.008$, $p > 0.99$. Consequently, all further analyses utilizing this task were conducted using composite scores of the two conditions.

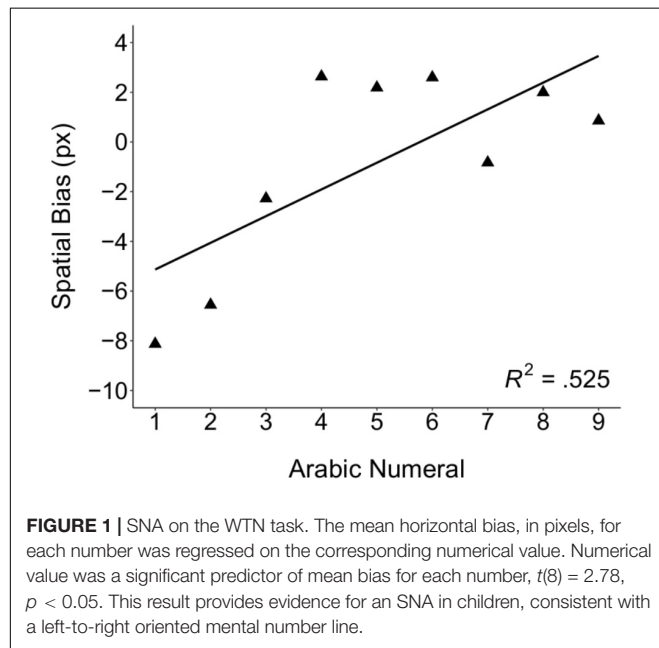
As a measure of SNAs on this task, we calculated a total congruency score for each child (see **Table 2**). Congruency scores were calculated as the difference between correct congruent and incongruent trials such that a positive congruency score represented a rightward oriented SNA effect. Children exhibited congruency scores ($M = 0.594$, $SD = 2.23$) that were significantly

³Analysis of children's bias along the vertical axis yielded no evidence of a significant SNA, $p > 0.70$. Previous research suggests SNAs along the vertical axis may be less robust than those along the horizontal axis (Holmes and Lourenco, 2012). However, it is also possible that the rectangular space used here constrained bias in the vertical dimension.

TABLE 2 | Partial correlations between SNA tasks and the different math tasks (dependent variable for each task in parentheses), controlling for age. Also included are descriptive statistics for each task.

Task	1	2	3	4	5	<i>n</i>	<i>M</i>	<i>SD</i>	Reliability
1 WTN (slope)		0.012	0.316	0.100	0.747	58	1.07	4.06	0.71
2 Magnitude comparison (congruency score)	0.339* [†]		0.021	0.770	0.153	64	0.594	2.23	0.52
3 ACA (proportion correct)	-0.148	0.313*		0.356	0.690	56	0.659	0.114	0.72
4 ASA (proportion correct)	0.248	-0.042	0.138		0.055	53	0.755	0.148	0.86
5 WJ–Calculation (raw scores)	0.044	-0.182	0.055	0.267		54	6.71	4.21	0.93 [^]

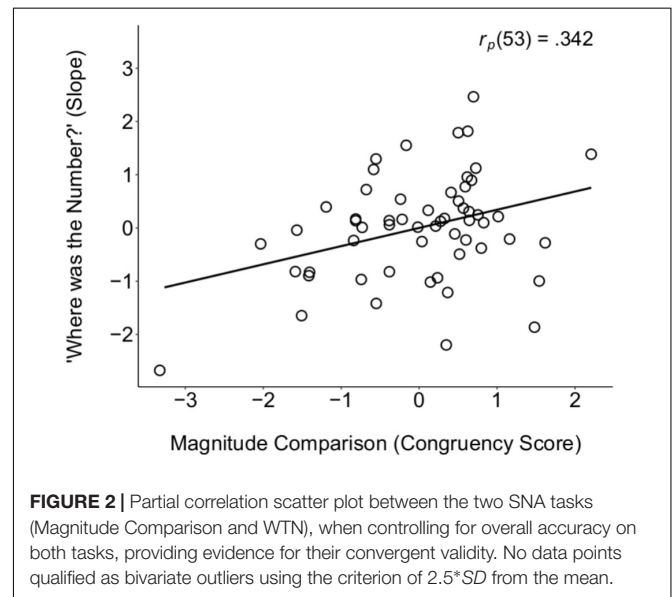
Values below the diagonal represent Pearson correlation coefficients (*r*). Values above the diagonal are *p*-values (uncorrected). Values in the last four columns display descriptive data for all tasks. **p* < 0.05. [†]Partial correlation additionally controlling for accuracy on both tasks. [^]Reported reliability based on a split-half procedure (McGrew et al., 2007).



greater than zero, $t(63) = 2.13$, $p = 0.037$, $d = 0.266$, suggesting a left-to-right mental number line that applies to non-symbolic displays of number. Although not a significant majority of children (binomial test, $p = 0.191$), more than half of them displayed positive congruency scores (36 of 64; 56%). Children's congruency scores were not significantly correlated with age, $r(62) = -0.210$, $p = 0.095$, suggesting no relation between non-symbolic SNAs and age in a sample of 5- to 7-year-old children. Children's congruency scores were significantly negatively correlated with overall accuracy, $r(62) = -0.250$, $p = 0.047$, but this relation was no longer significant after controlling for age, $r_p(61) = -0.191$, $p = 0.134$, suggesting no specific relation between congruency scores and accuracy beyond that accounted for by age.

Relations Between SNA Tasks

To test for a potential relation between the two SNA tasks (WTN and magnitude comparison), we conducted correlation analyses between children's slopes on the WTN task and their congruency scores on the magnitude comparison task. When controlling for accuracy on the two tasks to account for differences in task



demands, there was a significant correlation between children's performance on the two SNA tasks, $r_p(53) = 0.342$, $p = 0.011$ (see **Figure 2**), demonstrating convergent validity for these SNAs, and suggesting a left-to-right mental number line that is robust to the type of stimuli (symbolic and non-symbolic number). The relation between these two tasks held when additionally controlling for age, $r_p(52) = 0.339$, $p = 0.012$, suggesting further that the left-to-right mental number line is stable within the age range tested (5 to 7 years). Moreover, in addition to age, this relation held when further controlling for general cognitive abilities – namely, verbal proficiency (WJ–Picture Vocabulary), working memory (WJ–Auditory Working Memory), and short-term memory (K–ABC Spatial Memory), $r_p(49) = 0.302$, $p = 0.034$ (analyses conducted on raw scores of each task, discussed further below).

Children's Performance on the Math Tasks

ACA Task

Nine children were excluded from the analyses of the ACA task for failing to complete one or both conditions of this task (see **Table 2** for descriptive data). In the remaining sample

($n = 56$), performance (proportion correct) was significantly above the chance level of 0.50 ($M = 0.659$, $SD = 0.114$), $t(55) = 10.50$, $p < 0.001$, $d = 1.40$. A repeated measures analysis of variance (ANOVA) with operation (addition and subtraction) and ratio (4:5, 4:6, and 4:7) as the within-subjects factors revealed a marginal effect of operation, $F(1, 55) = 3.81$, $p = 0.056$, $\eta_p^2 = 0.065$, such that children performed somewhat better on addition than subtraction trials (Barth et al., 2008). There was also a significant effect of ratio, $F(2, 110) = 7.11$, $p = 0.001$, $\eta_p^2 = 0.115$, and a linear contrast analysis revealed that performance improved as ratio decreased (e.g., better performance for a ratio of 4:7 than 4:5), $F(1, 55) = 11.64$, $p = 0.001$, $\eta_p^2 = 0.175$, as would be expected if the computations were performed approximately (Lipton and Spelke, 2004; Halberda and Feigenson, 2008). There was no interaction between operation and ratio, $p > 0.94$.

ASA Task

Thirteen children were excluded from the analyses of the ASA task for failing to complete one or both conditions of this task (see Table 2). In the remaining sample ($n = 53$), performance (proportion correct) was significantly above the chance level of 0.50 ($M = 0.755$, $SD = 0.148$), $t(52) = 12.56$, $p < 0.001$, $d = 1.73$. A repeated measures ANOVA with operation (addition and subtraction) and ratio (4:5, 4:6, and 4:7) as within-subjects factors revealed a significant effect of operation, $F(1, 52) = 4.66$, $p = 0.035$, $\eta_p^2 = 0.082$, such that children were more accurate on addition than subtraction trials. There was a marginally significant effect of ratio, $F(2, 104) = 3.02$, $p = 0.053$, $\eta_p^2 = 0.055$, and a linear contrast analysis revealed a statistically significant linear trend, $F(1, 55) = 4.45$, $p = 0.040$, $\eta_p^2 = 0.079$, such that performance improved as ratio decreased (e.g., better performance for a ratio of 4:7 than 4:5), as expected. There was no interaction between operation and ratio, $p > 0.13$.

Exact Symbolic Arithmetic Task

Although the WJ–Calculation subtest allows for computing standardized scores, we instead utilized raw scores in our analyses, as in other studies (Bugden and Ansari, 2011; Lourenco and Bonny, 2017). The use of raw scores allowed for the inclusion of all children in the subsequent correlation analyses because standardized scores could not be calculated for several children who received scores of zero ($n = 9$). Raw scores on WJ–Calculation ranged from 0 to 14 ($M = 6.71$, $SD = 4.21$). As indicated above, these scores were normally distributed.

Analyses of the Relations Between Math Tasks

To our knowledge, previous studies have not examined the relations between the math tasks used in the present study. To assess the potential relations between these tasks, we first conducted a series of partial correlations, controlling for age (see Table 2 for correlations). Despite acceptable reliability for each measure ($r_s > 0.52$), we did not observe significant correlations among the math measures, with the exception of one marginal trend in the relation between tasks that shared a common symbolic format, ASA and WJ–Calculation, $r_p(50) = 0.267$, $p = 0.055$, such that children who performed better on the ASA task also tended to perform better on

WJ–Calculation. These findings are consistent with the literature on math abilities in adults in which dissociations between abilities within the math domain have been reported (e.g., Rosenberg-Lee et al., 2011; Lourenco et al., 2012). Likewise, other work has shown that math abilities are, to some extent, dissociable at younger ages (Fuchs et al., 2010; LeFevre et al., 2010; Cho et al., 2011). Specifically, these dissociations may reflect differences in calculation type (approximate vs. exact), numerical format (non-symbolic vs. symbolic), modality (uni-modal vs. cross-modal), and/or presentation format (simultaneous vs. sequential). Although it is possible that the lack of significant relations observed in the present study could reflect attenuation due to task reliabilities, all reliabilities were in the acceptable range. Therefore, these findings likely reflect early developmental dissociations across different math tasks.

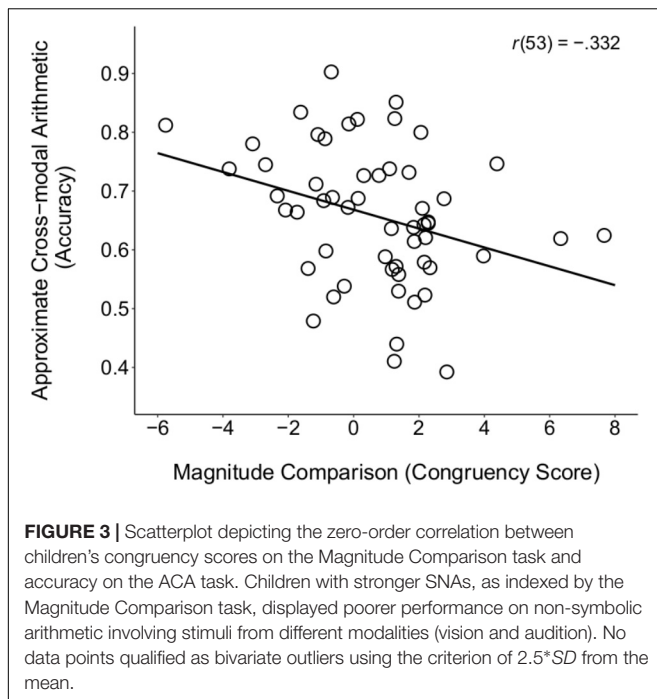
Control Tasks

Given that raw scores were used for the WJ–Calculation task, we likewise used raw scores for all of the control tasks. Scores on WJ–Picture Vocabulary, our measure of verbal proficiency, ranged from 14 to 27, with a mean of 20.55 ($SD = 3.05$). Scores on WJ–Auditory Working Memory, our measure of verbal working memory, ranged from 0 to 25, with a mean of 13.32 ($SD = 6.09$). Scores on K-ABC Spatial Memory, our measure of spatial short-term memory, ranged from 5 to 16, with a mean of 11.11 ($SD = 3.10$). As indicated above, these scores were normally distributed.

Is There a Relation Between Children's SNAs and Math Performance?

We conducted correlation analyses between children's performance on the two SNA tasks and each math task, controlling for age, to address the main question motivating the present work. When the WTN task served as the SNA measure, we found no significant correlations between children's slopes on the WTN task and their accuracy on any math task (see Table 2). In particular, there were no relations between children's performance on WTN and ACA tasks. Furthermore, there were no relations between children's performance on the WTN task and either symbolic arithmetic task (i.e., ASA and WJ–Calculation).

When using congruency scores on the magnitude comparison task as the measure of SNAs, we found a significant correlation with performance on the ACA task, $r_p(52) = -0.313$, $p = 0.021$ (see Figure 3; $p_s > 0.15$ for all other correlations between the magnitude comparison task and math ability, see Table 2). This negative correlation suggests that a stronger SNA was related to poorer understanding of cross-modal number representations that required arithmetic operations. Moreover, this effect held when additionally controlling for children's verbal proficiency (WJ–Picture Vocabulary), working memory (WJ–Auditory Working Memory), and short-term memory (K-ABC Spatial Memory), $r_p(49) = -0.314$, $p = 0.025$, suggesting a robust relation not due to these particular cognitive abilities. But could poor numerical precision (Halberda and Feigenson, 2008) and, thus, difficulty distinguishing smaller and larger numerical arrays, account for the significant correlation? We



addressed this possibility directly by controlling for children's accuracy on the magnitude comparison task, in addition to age and general cognitive ability. The relation between children's SNAs, as indexed by congruency on the magnitude comparison task, and ACA performance, remained statistically significant, $r_p(48) = -0.290$, $p = 0.041$. Thus, although there was only one significant correlation between children's SNAs and their math ability in the present study, this effect held when controlling for other cognitive abilities and when addressing an alternative account based on poor numerical precision. This finding suggests that there is a negative relation between the directional mental number line, as assessed by the magnitude comparison task, and the understanding of abstract (i.e., modality-independent) numerosity. We discuss this negative relation in the Section "General Discussion."

General Discussion

The primary goal of the present study was to examine the potential relations between SNAs and emerging mathematical competence in childhood. Although much interest has concerned the spatial nature of number representations, we know little about the links between these representations and mathematical development. As discussed in the Section "Introduction," existing research on this topic has been mixed (e.g., Hoffman et al., 2013; Gibson and Maurer, 2016). Here we adopted two measures of SNAs and multiple measures of math competence in an effort to shed light on the important question of whether the directionality of the mental number line may offer functional significance in the domain of mathematics, particularly at an age when quantitative reasoning is undergoing development.

Our two measures of SNAs revealed left-to-right orientation of number representations in 5- to 7-year-olds. We showed this

effect with a non-symbolic magnitude comparison task, which has been used in previous work with children (Patro and Haman, 2012), as well as the novel, symbolic WTN task, only used previously with adults (Aulet et al., 2017). Importantly, not only did we find evidence of left-to-right orientation of number on both tasks, but we also found a correlation between performance on these tasks, even when controlling for accuracy, age, and general cognitive abilities, thereby providing convergent evidence of a mental number line early in development. Even in adults, it is rare to assess construct validity of SNAs (for exceptions, see Cheung et al., 2015; Georges et al., 2017a). Here, we show that SNAs can be captured with different tasks in children, such that individual differences in the strength of these SNAs were common across tasks.

We also examined the relation between each SNA task and children's performance on a variety of measures designed to tap basic mathematical competence. We observed no significant correlations between slopes on the WTN task and children's performance on the math tasks, suggesting no relation between SNAs and early math abilities. However, could other factors account for the lack of correlations between the WTN task and math performance? One possibility is that slopes on this task underestimated the directionality of the mental number line for children with more compressive mental number lines. Visual inspection of **Figure 1** certainly suggests a non-linear relation between numerical value and spatial bias, which may reflect compressive representations of number on this task. As the goal of the present study was to assess the relation between directionality and math ability, we did not systematically investigate spatial scaling of the mental number line on the WTN task. Nonetheless, although we cannot rule out this possibility directly, we think it is unlikely that slopes were systematically underestimated given that the majority of children's responses were consistent with a rightward-oriented mental number line. Moreover, we observed a significant positive correlation between children's slopes on the WTN task and congruency scores on the magnitude comparison task, which would not be expected if the underestimation of slopes on the WTN task resulted in a failure to capture individually differences in the directionality of children's mental number lines. Thus, although it is possible that individual differences in spatial scaling may have impacted the precision of the estimates of directionality on the WTN task, this alone likely cannot account for the non-existent relations between WTN slopes and mathematical ability.

By contrast, there was a relation between children's congruency scores on the magnitude comparison task and their performance on the ACA task, but this relation was negative, which we did not predict for children between 5 and 7 years of age. In particular, we found that 5- to 7-year-olds with stronger SNAs (i.e., larger congruency scores) performed worse on the ACA task, even after controlling for age, general cognitive abilities, such as working memory, and accuracy on the magnitude comparison task itself. As discussed in the Section "Introduction," although previous studies in children have typically reported a positive relation between SNA strength and math ability (Bachot et al., 2005; Georges et al., 2017b), our findings mirror previous studies in adults that

have also reported a negative relation between SNA strength and mathematical ability (Hoffmann et al., 2014a; Cipora et al., 2016). At minimum, the data observed in the present study would seem to suggest that children with a more robust left-to-right mental number line perform at a level below their peers in mathematics. This negative relation could be interpreted as suggesting that a directional mental number line hinders, rather than facilitates, mathematical development. Such an interpretation, however, is at odds with a large literature on the role of analogy (Gentner et al., 2001; Siegler, 2016), metaphor (Núñez and Lakoff, 2005), and embodiment (Barsalou, 2008) in the acquisition and understanding of abstract concepts.

One possible explanation for the negative relation between children's math performance and the magnitude comparison task, but not the WTN task, is that this relation may reflect additional task-specific demands. In particular, the congruency effect on the magnitude comparison task, which was used to assess the strength of children's SNAs, might reflect inhibitory control required by this specific task and potentially associated with mathematical competence (Fuhs and McNeil, 2013; Cragg and Gilmore, 2014; Hohol et al., 2017). Successful performance on the magnitude comparison task required an assessment of which array was smaller or larger in numerosity, regardless of the spatial position of the arrays. On the incongruent trials, this might involve inhibition of the mental number line, since, on these trials, the correct array was in the spatially incongruent position. As a consequence, inhibition of the mental number line would actually result in greater accuracy on these incongruent trials. Thus, smaller congruency scores could indicate a weak SNA or could, instead, indicate an inability to inhibit an SNA when it conflicted with the goal of the task. By contrast, the WTN task required no such inhibitory demands and, as discussed earlier, this task was not correlated with any of the math measures given to children.

The ACA task, like the magnitude comparison task, displayed arrays on the left and right sides of the screen. Could this common spatial layout therefore explain the negative relation between congruency scores on the magnitude comparison task and accuracy on the ACA task? Although we cannot rule out this possibility directly, we would suggest that it is unlikely because the ASA task also shared this layout, and there was no relation between congruency scores on the magnitude comparisons task and accuracy on the ASA task. Thus, the common layout between tasks would appear insufficient to explain the negative relation observed between SNAs and math ability in the present study. What, then, might account for this finding? Successful performance on the ACA task might also depend on inhibitory control, similar to the suggestion by Fuhs et al. (2016) that a relation between executive function and math achievement arises from the ability to accurately represent the value of a numerical set as opposed to the individual items within a set. In the ACA task, numerosities were presented across different modalities (vision and audition) and presentation formats (simultaneous vs. sequential). On this task, in contrast to the ASA task, children had to abstract numerical value over quite disparate stimuli. The differences in modality and presentation format might have

increased the salience of the individual items, requiring more inhibition to delay responses until the value of the full set could be assessed. If inhibitory control were necessary to assess the set as a whole, then individual differences in inhibitory control would influence performance on the ACA task. As a consequence, poor inhibitory control could lead to both larger congruency scores on the magnitude comparison task and worse performance on the ACA task (for a similar finding, see Hoffmann et al., 2014b).

Given the alternative explanation just described, and the lack of significant correlations involving the WTN task (our other measure of SNAs), our findings do not provide strong support for a relation between a directional mental number line and math competence in 5- to 7-year-old children. Importantly, the inhibitory control account of the negative relation between SNAs, as indexed by the magnitude comparison task, and accuracy on the ACA task, does not suggest that the mental number line itself is negatively related to math ability. Rather, tasks such as the magnitude comparison task may require inhibition of the mental number line for optimal performance and other tasks may depend on inhibitory control more generally for performing numerical comparison and/or arithmetic computation across numerical format (Fuhs et al., 2016). Thus, it remains possible that, in the absence of such inhibitory control demands, there may exist a positive relation between SNAs involving non-symbolic numerosities and math ability.

As we outlined in the Section "Introduction," Hoffman et al. (2013) found a positive correlation between SNAs and math ability. Interestingly, they used a magnitude comparison task, as in the present study, but with numerals. This study, however, did not find an overall effect of SNAs for the group of children tested, nor were there controls for general cognitive functioning, which could account for a correlation between performance on their magnitude comparison task and numerical proficiency. As in the present study, it would be especially important to determine the extent to which inhibitory control might account for the correlation in Hoffman et al. (2013). Other published work has found no significant effects between the strength of children's SNAs and performance on a math test (Gibson and Maurer, 2016). This study also did not include measures of general cognitive functioning, such that it is unclear to what extent inhibitory control or other variables could have accounted for the results. It is also possible that age may play a role in determining the relation between SNAs and math ability. An important difference between these studies is that the children in the Hoffman et al. (2013) study were younger than those tested here and in the study of Gibson and Maurer (2016). Thus, a positive SNA-math link could exist earlier in development, such that a mental number line might prove beneficial to mathematical reasoning, but this link is only present during the earliest stages of acquisition when young children are first learning quantitative concepts and operations such as those tested in the present study.

An important consideration for this research program going forward is whether an individual differences approach, adopted here and in other studies, is well suited for assessing whether a mental number line benefits math development. In particular,

we took the approach that if the directionality of the mental number line were relevant for math development, then one should observe a correlation between the strength of one's SNA and performance on one or more of the math tasks administered to children. However, it is possible that some minimal amount of left-to-right organization in one's number representations is sufficient for supporting learning of abstract number concepts or performing arithmetic computations. If minimal organization were sufficient, then relations between tasks assessing SNAs and children's math performance would not be observed.

Another important consideration that follows from the current and existing research is that other components of the mental number line, besides directionality, may be related to mathematical competence (for review, see Cipora et al., 2015). In the Section "Introduction," we hypothesized that the spatial grounding provided by a mental number line might facilitate understanding of number as an abstract concept and, thus, a stronger left-to-right orientation of number would provide support for math tasks, such as cross-modal arithmetic, that rely on this abstract understanding (Lakoff and Núñez, 2000; Barsalou, 2008). We also hypothesized that directionality could be beneficial for performing arithmetic. Effects of operational momentum in which individuals associate larger outcomes with addition and smaller outcomes with subtraction (McCrink et al., 2007; Knops et al., 2009) are consistent with shifts of attention along a mental number line during these arithmetic operations. The ability to dynamically shift one's attention in relation to this spatial representation may be comparable to other visuospatial processes such as mental rotation that have been shown to relate to mathematical reasoning (Thompson et al., 2013; Cheng and Mix, 2014). The magnitude comparison and WTN tasks, however, were designed to capture the extent of left-to-right orientation, not the dynamic quality of the mental number line, or of attentional processes that may be applied to it, which, ultimately, may be more predictive of math development.

Another critical feature of the mental number line is the spatial scaling of numerical intervals. Rather than direction (e.g., left-to-right), we can ask whether the scaling is best characterized by a linear or logarithmic mapping of number to spatial extent. Most commonly, these mappings are measured by a number line estimation task where participants designate the position of a numerical value on a physical line anchored by two numbers (Siegler and Opfer, 2003). In the case of a linear representation, a change in numerical distance corresponds to an equivalent change in spatial distance. That is, across the entire range of the number line, numerical values are represented with consistent spatial intervals when the representation is linear. Conversely, for compressive representations, the spatial distance between two small numbers is judged as larger than that between two larger numbers of equivalent numerical difference (e.g., children designate the numbers 5 and 15 as farther apart than 75 and 85; but, see Barth et al., 2011; Cohen and Quinlan, 2017). Not only has the linearity of one's number line been shown to correlate positively with math proficiency, as measured by a variety of math measures, but causal evidence

has also been put forth, in which children who receive training to increase the linearity of their numerical representations subsequently show better math scores than those receiving non-numerical (control) training (for meta-analysis, see Schneider et al., 2018). Thus, although there may be little evidence for a relation between the directionality of the mental number line and mathematical competence, there is accumulating support for the importance of a linear mental number line in math development.

CONCLUSION

In conclusion, our results do not provide strong support for a relation between a directional mental number line and mathematical ability in 5- to 7-year-old children. Contrary to our initial prediction, the sole significant SNA-math relation was negative, such that a stronger SNA was associated with worse, not better, performance on a measure of early mathematical competence. Though we have suggested that this link is likely due to individual differences in inhibitory control, consistent with previous research (Hoffmann et al., 2014b), we acknowledge the speculative nature of this claim given that the present study did not include a direct measure of inhibition. Thus, we urge future research on this topic to consider the potential influence of inhibitory control on different measures. Moreover, additional research is necessary to determine whether children younger than those tested here are more likely to benefit from a left-to-right oriented mental number line, and further, whether different facets of the mental number line, such as directionality and scale, contribute differentially to mathematical development. We also encourage researchers to consider experimental designs beyond an individual differences approach to shed light on these questions.

ETHICS STATEMENT

This study was carried out in accordance with the guidelines of the Institutional Review Board (IRB) at Emory University under IRB protocol #003571. The protocol was approved by the IRB at Emory University. The parent or guardian of all participants gave written informed consent in accordance with the Declaration of Helsinki.

AUTHOR CONTRIBUTIONS

LA and SL conceived and designed the experiments, and wrote the paper. LA performed the experiments and analyzed the data.

FUNDING

This work was supported by a National Institutes of Health (NIH) institutional training grant (T32 HD071845) to LA and a Scholar Award from the John Merck Fund to SL. All views expressed are solely those of the authors.

REFERENCES

- Abrahamse, E., van Dijck, J. P., and Fias, W. (2016). How does working memory enable number-induced spatial biases? *Front. Psychol.* 7:977. doi: 10.3389/fpsyg.2016.00977
- Agrillo, C., Piffer, L., and Bisazza, A. (2011). Number versus continuous quantity in numerosity judgments by fish. *Cognition* 119, 281–287. doi: 10.1016/j.cognition.2010.10.022
- Anobile, G., Castaldi, E., Turi, M., Tinelli, F., and Burr, D. C. (2016). Numerosity but not texture-density discrimination correlates with math ability in children. *Dev. Psychol.* 52, 1206–1216. doi: 10.1037/dev0000155
- Aulet, L. S., Yousif, S. R., and Lourenco, S. F. (2017). “Numbers uniquely bias spatial attention: a novel paradigm for understanding spatial-numerical associations,” in *Proceedings of the 39th Annual Conference of the Cognitive Science Society*, eds G. Gunzelmann, A. Howes, T. Tenbrink, and E. J. Davelaar (Austin, TX: Cognitive Science Society), 75–80.
- Bachot, J., Gevers, W., Fias, W., and Roeyers, H. (2005). Number sense in children with visuospatial disabilities: orientation of the mental number line. *Psychol. Sci.* 47, 172–183.
- Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.* 59, 617–645. doi: 10.1146/annurev.psych.59.103006.093639
- Barth, H., Beckmann, L., and Spelke, E. (2008). Nonsymbolic, approximate arithmetic in children: evidence for abstract addition prior to instruction. *Dev. Psychol.* 44, 1466–1477. doi: 10.1037/a0013046
- Barth, H., Slusser, E., Cohen, D., and Paladino, A. M. (2011). A sense of proportion: commentary on Opfer, Siegler, & Young. *Dev. Sci.* 14, 1205–1206. doi: 10.1111/j.1467-7687.2011.01081.x
- Bonny, J. W., and Lourenco, S. F. (2013). The approximate number system and its relation to early math achievement: evidence from the preschool years. *J. Exp. Child Psychol.* 114, 375–388. doi: 10.1016/j.jecp.2012.09.015
- Booth, J. L., and Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Dev. Psychol.* 42, 189–201. doi: 10.1037/0012-1649.41.6.189
- Booth, J. L., and Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Dev.* 79, 1016–1031. doi: 10.1111/j.1467-8624.2008.01173.x
- Bugden, S., and Ansari, D. (2011). Individual differences in children’s mathematical competence are related to the intentional but not automatic processing of Arabic numerals. *Cognition* 118, 32–44. doi: 10.1016/j.cognition.2010.09.005
- Bulf, H., de Hevia, M. D., and Macchi Cassia, V. (2016). Small on the left, large on the right: numbers orient visual attention onto space in preverbal infants. *Dev. Sci.* 19, 394–401. doi: 10.1111/desc.12315
- Bull, R., Cleland, A. A., and Mitchell, T. (2013). Sex differences in the spatial representation of number. *J. Exp. Psychol. Gen.* 142, 181–192. doi: 10.1037/a0028387
- Cantlon, J. F., and Brannon, E. M. (2006). Shared system for ordering small and large numbers in monkeys and humans. *Psychol. Sci.* 17, 401–406. doi: 10.1111/j.1467-9280.2006.01719.x
- Cheng, Y. L., and Mix, K. S. (2014). Spatial training improves children’s mathematics ability. *J. Cogn. Dev.* 15, 2–11. doi: 10.1080/15248372.2012.725186
- Cheung, C.-N., Ayzenberg, V., Diamond, R. F. L., Yousif, S. R., and Lourenco, S. F. (2015). “Probing the mental number line: a between-task analysis of spatial-numerical associations,” in *Proceedings of the 37th Annual Conference of the Cognitive Science Society*, (Hillsdale, NJ: Lawrence Erlbaum Associates), 357–362.
- Cho, S., Ryali, S., Geary, D. C., and Menon, V. (2011). How does a child solve 7 + 8? Decoding brain activity patterns associated with counting and retrieval strategies. *Dev. Sci.* 14, 989–1001. doi: 10.1111/j.1467-7687.2011.01055.x
- Cipora, K., Hohol, M., Nuerk, H.-C., Willmes, K., Brożek, B., Kucharzyk, B., et al. (2016). Professional mathematicians differ from controls in their spatial-numerical associations. *Psychol. Res.* 80, 710–726. doi: 10.1007/s00426-015-0677-6
- Cipora, K., and Nuerk, H.-C. (2013). Is the SNARC effect related to the level of mathematics? No systematic relationship observed despite more power, more repetitions, and more direct assessment of arithmetic skill. *Q. J. Exp. Psychol.* 66, 1974–1991. doi: 10.1080/17470218.2013.772215
- Cipora, K., Patro, K., and Nuerk, H.-C. (2015). Are spatial-numerical associations a cornerstone for arithmetic learning? The lack of genuine correlations suggests no. *Mind Brain Educ.* 9, 190–206. doi: 10.1111/mbe.12093
- Cohen, D. J., and Quinlan, P. T. (2017). The log-linear response function of the bounded number-line task is unrelated to the psychological representation of quantity. *Psychon. Bull. Rev.* 25, 447–454. doi: 10.3758/s13423-017-1290-z
- Cooperrider, K., Marghetis, T., and Núñez, R. (2017). Where does the ordered line come from? Evidence from a culture of Papua New Guinea. *Psychol. Sci.* 28, 599–608. doi: 10.1177/0956797617691548
- Cordes, S., and Brannon, E. M. (2008). Quantitative competencies in infancy. *Dev. Sci.* 11, 803–808. doi: 10.1111/j.1467-7687.2008.00770.x
- Cragg, L., and Gilmore, C. (2014). Skills underlying mathematics: the role of executive function in the development of mathematics proficiency. *Trends Neurosci. Educ.* 3, 63–68. doi: 10.1016/j.tine.2013.12.001
- de Hevia, M. D., Addabbo, M., Girelli, L., and Macchi Cassia, V. (2014). Human infants’ preference for left-to-right oriented increasing numerical sequences. *PLoS One* 9:e96412. doi: 10.1371/journal.pone.0096412
- Dehaene, S., Bossini, S., and Giraux, P. (1993). The mental representation of parity and number magnitude. *J. Exp. Psychol. Gen.* 122, 371–396. doi: 10.1037/0096-3445.122.3.371
- Dehaene, S., Izard, V., Spelke, E., and Pica, P. (2008). Log or linear? Distinct intuitions of the number scale in western and Amazonian indigene cultures. *Science* 320, 1217–1220. doi: 10.1126/science.1156540
- Drucker, C., and Brannon, E. M. (2014). Rhesus monkeys (*Macaca mulatta*) map number onto space. *Cognition* 132, 57–67. doi: 10.1016/j.cognition.2014.03.011
- Feigenson, L., Dehaene, S., and Spelke, E. (2004). Core systems of number. *Trends Cogn. Sci.* 8, 307–314. doi: 10.1016/j.tics.2004.05.002
- Fischer, M. H., Castel, A. D., Dodd, M. D., and Pratt, J. (2003). Perceiving numbers causes spatial shifts of attention. *Nat. Neurosci.* 6, 555–556. doi: 10.1038/nn1066
- Fischer, M. H., and Shaki, S. (2014). Spatial associations in numerical cognition – From single digits to arithmetic. *Q. J. Exp. Psychol.* 67, 1461–1483. doi: 10.1080/17470218.2014.927515
- Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., Seethaler, P. M., et al. (2010). Do different types of school mathematics development depend on different constellations of numerical versus general cognitive abilities? *Dev. Psychol.* 46, 1731–1746. doi: 10.1037/a0020662
- Fuhs, M. W., Hornburg, C. B., and McNeil, N. M. (2016). Specific early number skills mediate the association between executive functioning skills and mathematics achievement. *Dev. Psychol.* 52, 1217–1235. doi: 10.1037/dev0000145
- Fuhs, M. W., and McNeil, N. M. (2013). ANS acuity and mathematics ability in preschoolers from low-income homes: contributions of inhibitory control. *Dev. Sci.* 16, 136–148. doi: 10.1111/desc.12013
- Galton, F. (1880). Visualised numerals. *Nature* 21, 252–256. doi: 10.1038/021252a0
- Gazes, R. P., Diamond, R. F., Hope, J. M., Caillaud, D., Stoinski, T. S., and Hampton, R. R. (2017). Spatial representation of magnitude in gorillas and orangutans. *Cognition* 168, 312–319. doi: 10.1016/j.cognition.2017.07.010
- Gebuis, T., and Reynvoet, B. (2011). Generating non-symbolic number stimuli. *Behav. Res. Methods* 43, 981–986. doi: 10.3758/s13428-011-0097-5
- Gentner, D., Holyoak, K. J., and Kokinov, B. N. (2001). *The Analogical Mind*. Cambridge, MA: MIT Press.
- Georges, C., Hoffmann, D., and Schiltz, C. (2017a). How and why do number-space associations co-vary in implicit and explicit magnitude processing tasks? *J. Numer. Cogn.* 3, 182–211.
- Georges, C., Hoffmann, D., and Schiltz, C. (2017b). Mathematical abilities in elementary school: do they relate to number-space associations? *J. Exp. Child Psychol.* 161, 126–147. doi: 10.1016/j.jecp.2017.04.011
- Gevers, W., Reynvoet, B., and Fias, W. (2003). The mental representation of ordinal sequences is spatially organized. *Cognition* 87, B87–B95. doi: 10.1016/S0010-0277(02)00234-2
- Gibson, L. C., and Maurer, D. (2016). Development of SNARC and distance effects and their relation to mathematical and visuospatial abilities. *J. Exp. Child Psychol.* 150, 301–313. doi: 10.1016/j.jecp.2016.05.009
- Gilmore, C. K., McCarthy, S. E., and Spelke, E. S. (2007). Symbolic arithmetic knowledge without instruction. *Nature* 447, 589–591. doi: 10.1038/nature05850
- Halberda, J., and Feigenson, L. (2008). Developmental change in the acuity of the “number sense”: the approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Dev. Psychol.* 44, 1457–1465. doi: 10.1037/a0012682

- Halberda, J., Mazocco, M. M., and Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature* 455, 665–670. doi: 10.1038/nature07246
- Hoffman, D., Hornung, C., Martin, R., and Schiltz, C. (2013). Developing number-space associations: SNARC effects using a color discrimination task in 5-year-olds. *J. Exp. Child Psychol.* 116, 775–791. doi: 10.1016/j.jecp.2013.07.013
- Hoffmann, D., Mussolin, C., Martin, R., and Schiltz, C. (2014a). The impact of mathematical proficiency on the number-space association. *PLoS One* 9:e85048. doi: 10.1371/journal.pone.0085048
- Hoffmann, D., Pigat, D., and Schiltz, C. (2014b). The impact of inhibition capacities and age on number–space associations. *Cogn. Process.* 15, 329–342. doi: 10.1007/s10339-014-0601-9
- Hohol, M., Cipora, K., Willmes, K., and Nuerk, H. C. (2017). Bringing back the balance: domain-general processes are also important in numerical cognition. *Front. Psychol.* 8:499. doi: 10.3389/fpsyg.2017.00499
- Holmes, K. J., Ayzenberg, V., and Lourenco, S. F. (2016). Gamble on gaze: eye movements reflect the numerical value of blackjack hands. *Psychon. Bull. Rev.* 23, 1974–1981. doi: 10.3758/s13423-016-1055-0
- Holmes, K. J., and Lourenco, S. F. (2012). Orienting numbers in mental space: horizontal organization trumps vertical. *Q. J. Exp. Psychol.* 65, 1044–1051. doi: 10.1080/17470218.2012.685079
- Kaufman, A. S., and Kaufman, N. L. (1983). *Kaufman Assessment Battery for Children*. Circle Pines, MN: American Guidance Service.
- Klein, E., Huber, S., Nuerk, H.-C., and Moeller, K. (2014). Operational momentum affects eye fixation behaviour. *Q. J. Exp. Psychol.* 67, 1614–1625. doi: 10.1080/17470218.2014.902976
- Knops, A., Viarouge, A., and Dehaene, S. (2009). Dynamic representations underlying symbolic and nonsymbolic calculation: evidence from the operational momentum effect. *Atten. Percept. Psychophys.* 71, 803–821. doi: 10.3758/APP.71.4.803
- Lakoff, G., and Núñez, R. E. (2000). *Where Mathematics Comes from: How the Embodied Mind Brings Mathematics into Being*. New York, NY: Basic books.
- Lauer, J. E., and Lourenco, S. F. (2016). Spatial processing in infancy predicts both spatial and mathematical aptitude in childhood. *Psychol. Sci.* 27, 1291–1298. doi: 10.1177/0956797616655977
- LeFevre, J. A., Fast, L., Skwarchuk, S. L., Smith-Chant, B. L., Bisanz, J., Kamawar, D., et al. (2010). Pathways to mathematics: longitudinal predictors of performance. *Child Dev.* 81, 1753–1767. doi: 10.1111/j.1467-8624.2010.01508.x
- Libertus, M. E., Feigenson, L., and Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Dev. Sci.* 14, 1292–1300. doi: 10.1111/j.1467-7687.2011.01080.x
- Lipton, J. S., and Spelke, E. S. (2004). Discrimination of large and small numerosities by human infants. *Infancy* 5, 271–290. doi: 10.1207/s15327078in0503_2
- Lourenco, S. F., and Bonny, J. W. (2017). Representations of numerical and non-numerical magnitude both contribute to mathematical competence in children. *Dev. Sci.* 20:e12418. doi: 10.1111/desc.12418
- Lourenco, S. F., Bonny, J. W., Fernandez, E. P., and Rao, S. (2012). Nonsymbolic number and cumulative area representations contribute shared and unique variance to symbolic math competence. *Proc. Natl. Acad. Sci. U.S.A.* 109, 18737–18742. doi: 10.1073/pnas.1207212109
- Lourenco, S. F., Cheung, C., and Aulet, L. S. (2018). “The predictive value of early visuospatial processing in mathematical development,” in *Heterogeneity of Function in Numerical Cognition*, eds A. Henik and W. Fias (San Diego, CA: Academic Press), 177–209. doi: 10.1037/a0025510
- Lourenco, S. F., and Longo, M. R. (2009). Multiple spatial representations of number: evidence for co-existing compressive and linear scales. *Exp. Brain Res.* 193, 151–156. doi: 10.1007/s00221-008-1698-9
- Mazocco, M. M., Feigenson, L., and Halberda, J. (2011). Preschoolers’ precision of the approximate number system predicts later school mathematics performance. *PLoS One* 6:e23749. doi: 10.1371/journal.pone.0023749
- McCrink, K., Dehaene, S., and Dehaene-Lambertz, G. (2007). Moving along the number line: operational momentum in nonsymbolic arithmetic. *Percept. Psychophys.* 69, 1324–1333. doi: 10.3758/BF03192949
- McCrink, K., Shaki, S., and Berkowitz, T. (2014). Culturally-driven biases in preschoolers’ spatial search strategies. *Cogn. Dev.* 30, 1–14. doi: 10.1016/j.cogdev.2013.11.002
- McGrew, K. S., Woodcock, R. W., and Schrank, K. A. (2007). *Woodcock-Johnson III Normative Update Technical Manual*. Rolling Meadows, IL: Riverside Pub.
- Núñez, R., and Lakoff, G. (2005). “The cognitive foundations of mathematics: the role of conceptual metaphor,” in *The Handbook of Mathematical Cognition*, ed. J. Campbell (New York, NY: Psychology Press), 109–124.
- Odic, D., Hock, H., and Halberda, J. (2014). Hysteresis affects number discrimination in young children. *J. Exp. Psychol. Gen.* 143, 255–265. doi: 10.1037/a0030825
- Opfer, J. E., and Furlong, E. E. (2011). How numbers bias preschoolers’ spatial search. *J. Cross Cult. Psychol.* 42, 682–695. doi: 10.1177/0022022111406098
- Opfer, J. E., Thompson, C. A., and Furlong, E. E. (2010). Early development of spatial numeric associations: evidence from spatial and quantitative performance of preschoolers. *Dev. Sci.* 13, 761–771. doi: 10.1111/j.1467-7687.2009.00934.x
- Opfer, J. E., Thompson, C. A., and Kim, D. (2016). Free versus anchored numerical estimation: a unified approach. *Cognition* 149, 11–17. doi: 10.1016/j.cognition.2015.11.015
- Patro, K., and Haman, M. (2012). The spatial–numerical congruity effect in preschoolers. *J. Exp. Child Psychol.* 111, 534–542. doi: 10.1016/j.jecp.2011.09.006
- Pinhas, M., and Fischer, M. H. (2008). Mental movements without magnitude? A study of spatial biases in symbolic arithmetic. *Cognition* 109, 408–415. doi: 10.1016/j.cognition.2008.09.003
- Rosenberg-Lee, M., Chang, T. T., Young, C. B., Wu, S., and Menon, V. (2011). Functional dissociations between four basic arithmetic operations in the human posterior parietal cortex: a cytoarchitectonic mapping study. *Neuropsychologia* 49, 2592–2608. doi: 10.1016/j.neuropsychologia.2011.04.035
- Rugani, R., Regolin, L., and Vallortigara, G. (2008). Discrimination of small numerosities in young chicks. *J. Exp. Psychol. Anim. Behav. Process.* 34, 388–399. doi: 10.1016/j.neuropsychologia.2011.04.035
- Rugani, R., Vallortigara, G., Priftis, K., and Regolin, L. (2015). Number-space mapping in the newborn chick resembles humans’ mental number line. *Science* 347, 534–536. doi: 10.1037/0097-7403.34.3.388
- Sasanguie, D., Göbel, S. M., Moll, K., Smets, K., and Reynvoet, B. (2013). Approximate number sense, symbolic number processing, or number–space mappings: what underlies mathematics achievement? *J. Exp. Child Psychol.* 114, 418–431. doi: 10.1016/j.jecp.2012.10.012
- Schneider, M., Grabner, R. H., and Paetsch, J. (2009). Mental number line, number line estimation, and mathematical achievement: their interrelations in Grades 5 and 6. *J. Educ. Psychol.* 101, 359–372. doi: 10.1016/j.jecp.2012.10.012
- Schneider, M., Merz, J., Sticker, J., De Smedt, B., Torbeyns, J., Verschaffel, L., et al. (2018). Associations of number line estimation with mathematical competence: a meta-analysis. *Child Dev.* doi: 10.1111/cdev.13068 [Epub ahead of print].
- Shaki, S., Fischer, M. H., and Göbel, S. M. (2012). Direction counts: a comparative study of spatially directional counting biases in cultures with different reading directions. *J. Exp. Child Psychol.* 112, 275–281. doi: 10.1016/j.jecp.2011.12.005
- Shaki, S., Fischer, M. H., and Petrusic, W. M. (2009). Reading habits for both words and numbers contribute to the SNARC effect. *Psychon. Bull. Rev.* 16, 328–331. doi: 10.3758/PBR.16.2.328
- Siegler, R. S. (2016). Magnitude knowledge: the common core of numerical development. *Dev. Sci.* 19, 341–361. doi: 10.1111/desc.12395
- Siegler, R. S., and Booth, J. L. (2004). Development of numerical estimation in young children. *Child Dev.* 75, 428–444. doi: 10.3758/PBR.16.2.328
- Siegler, R. S., and Opfer, J. E. (2003). The development of numerical estimation: evidence for multiple representations of numerical quantity. *Psychol. Sci.* 14, 237–243. doi: 10.1111/j.1467-8624.2004.00684.x
- Siegler, R. S., Thompson, C. A., and Opfer, J. E. (2009). The logarithmic-to-linear shift: one learning sequence, many tasks, many time scales. *Mind Brain Educ.* 3, 143–150. doi: 10.1037/a0014239
- Starr, A., Libertus, M. E., and Brannon, E. M. (2013). Number sense in infancy predicts mathematical abilities in childhood. *Proc. Natl. Acad. Sci. U.S.A.* 110, 18116–18120. doi: 10.1073/pnas.1302751110
- Tabachnick, B. G., and Fidell, L. S. (2001). *Using Multivariate Statistics*, 4th Edn. Boston, MA: Allyn & Bacon. doi: 10.1073/pnas.1302751110

- Thompson, J., Nuerk, H., Moeller, K., and Cohen Kadosh, R. (2013). The link between mental rotation ability and basic numerical representations. *Acta Psychol.* 144, 324–331. doi: 10.1016/j.actpsy.2013.05.009
- van Galen, M. S., and Reitsma, P. (2008). Developing access to number magnitude: a study of the SNARC effect in 7- to 9-year-olds. *J. Exp. Child Psychol.* 101, 99–113. doi: 10.1016/j.jecp.2008.05.001
- Verdine, B., Golinkoff, R. M., Hirsh-Pasek, K., and Newcombe, N. (2017). Links between spatial and mathematical skills across the preschool years. *Soc. Res. Child Dev. Monogr.* 82, 1–150.
- Wood, G., Willmes, K., Nuerk, H. C., and Fischer, M. H. (2008). On the cognitive link between space and number: a meta-analysis of the SNARC effect. *Psychol. Sci. Q.* 50, 489–525.
- Woodcock, R. W., McGrew, K. S., and Mather, N. (2001a). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside.
- Woodcock, R. W., McGrew, K. S., and Mather, N. (2001b). *Woodcock-Johnson III Tests of Cognitive Abilities*. Itasca, IL: Riverside.
- Xu, F., and Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition* 74, B1–B11. doi: 10.1016/S0010-0277(99)00066-9
- Zebian, S. (2005). Linkages between number concepts, spatial thinking, and directionality of writing: the SNARC effect and the reverse SNARC effect in English and Arabic monoliterates, biliterates, and illiterate Arabic speakers. *J. Cogn. Cult.* 5, 165–190. doi: 10.1163/1568537054068660
- Zorzi, M., Priftis, K., and Umiltà, C. (2002). Brain damage: neglect disrupts the mental number line. *Nature* 417, 138–139. doi: 10.1038/417138a

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Aulet and Lourenco. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Developmental Trajectory of the Operational Momentum Effect

Pedro Pinheiro-Chagas^{1,2*†}, Daniele Didino^{3*†}, Vitor G. Haase^{4,5,6,7}, Guilherme Wood^{8,9}
and André Knops^{3,10,11}

¹ Cognitive Neuroimaging Unit, CEA DRF/I2BM, INSERM, Université Paris-Sud, Université Paris-Saclay, NeuroSpin Center, Orsay, France, ² Laboratory of Behavioral and Cognitive Neuroscience, Stanford Human Intracranial Cognitive Electrophysiology Program, Department of Neurology and Neurological Sciences, Stanford University, Stanford, CA, United States, ³ Department of Psychology, Faculty of Life Sciences, Humboldt-Universität zu Berlin, Berlin, Germany, ⁴ Developmental Neuropsychology Laboratory (LND), Department of Psychology, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, ⁵ Programa de Pós-Graduação em Neurociências, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, ⁶ Department of Psychology, Graduate Program in Psychology, Cognition and Behavior – Graduate Program in Neuroscience, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, ⁷ Instituto Nacional de Ciência e Tecnologia sobre Comportamento, Cognição e Ensino, Universidade Federal de São Carlos, São Carlos, Brazil, ⁸ Department of Psychology, University of Graz, Graz, Austria, ⁹ BioTechMed-Graz, University of Graz, Graz, Austria, ¹⁰ CNRS UMR 8240, Laboratory for the Psychology of Child Development and Education, Paris, France, ¹¹ University Paris Descartes, Sorbonne Paris Cité, Paris, France

OPEN ACCESS

Edited by:

Maciej Haman,
University of Warsaw, Poland

Reviewed by:

Christine Schiltz,
University of Luxembourg,
Luxembourg
Mojtaba Soltanlou,
Eberhard Karls Universität Tübingen,
Germany

*Correspondence:

Daniele Didino
didiwoda@hu-berlin.de
Pedro Pinheiro-Chagas
ppinheirochagas@gmail.com

[†] These authors have contributed
equally to this work.

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 22 December 2017

Accepted: 06 June 2018

Published: 17 July 2018

Citation:

Pinheiro-Chagas P, Didino D,
Haase VG, Wood G and Knops A
(2018) The Developmental Trajectory
of the Operational Momentum Effect.
Front. Psychol. 9:1062.
doi: 10.3389/fpsyg.2018.01062

Mental calculation is thought to be tightly related to visuospatial abilities. One of the strongest evidence for this link is the widely replicated operational momentum (OM) effect: the tendency to overestimate the result of additions and to underestimate the result of subtractions. Although the OM effect has been found in both infants and adults, no study has directly investigated its developmental trajectory until now. However, to fully understand the cognitive mechanisms lying at the core of the OM effect it is important to investigate its developmental dynamics. In the present study, we investigated the development of the OM effect in a group of 162 children from 8 to 12 years old. Participants had to select among five response alternatives the correct result of approximate addition and subtraction problems. Response alternatives were simultaneously presented on the screen at different locations. While no effect was observed for the youngest age group, children aged 9 and older showed a clear OM effect. Interestingly, the OM effect monotonically increased with age. The increase of the OM effect was accompanied by an increase in overall accuracy. That is, while younger children made more and non-systematic errors, older children made less but systematic errors. This monotonous increase of the OM effect with age is not predicted by the compression account (i.e., linear calculation performed on a compressed code). The attentional shift account, however, provides a possible explanation of these results based on the functional relationship between visuospatial attention and mental calculation and on the influence of formal schooling. We propose that the acquisition of arithmetical skills could reinforce the systematic reliance on the spatial mental number line and attentional mechanisms that control the displacement along this metric. Our results provide a step in the understanding of the mechanisms underlying approximate calculation and an important empirical constraint for current accounts on the origin of the OM effect.

Keywords: operational momentum, approximate addition, approximate subtraction, children, development, attentional shift account, compression account, heuristic account

INTRODUCTION

Adults and children (Barth et al., 2006), and even infants (Wynn, 1992), are able to perform approximate mental calculation, which consists in the capacity to add or subtract numbers expressed in non-symbolic notations (e.g., dots). This skill requires to estimate the numerosity (i.e., cardinality) of two sets of elements and to encode it on an internal representation on which cognitive processes operate to generate the approximate outcome of the calculation. Growing evidence (McCrink et al., 2007; Pinhas and Fischer, 2008; Knops et al., 2009b; McCrink and Wynn, 2009; Lindemann and Tira, 2011; Chen and Verguts, 2012; Knops et al., 2013, 2014; Klein et al., 2014; Marghetis et al., 2014; Pinheiro-Chagas et al., 2017) shows that approximate addition and subtraction are subjected to an Operational Momentum (hereafter, OM) effect: results of addition are overestimated and results of subtraction are underestimated. Although an OM effect has been found in infants (McCrink and Wynn, 2009) and an inverse OM effect emerged in 6/7 years old children (Knops et al., 2013), no studies investigated the developmental trajectory of this effect. Therefore, it is still unclear how the OM effect evolves during the acquisition of formal mathematical knowledge. The relevance of the OM effect lies in the knowledge it provides regarding the cognitive mechanisms involved in the representation and the manipulation of non-symbolic numerical magnitudes. In this study, we aimed to measure how the OM effect evolves in children between 8 and 12 years of age. Moreover, the developmental trajectory of the OM effect can also provide evidence in favor of or against the current accounts proposed to explain this effect.

A prerequisite to perform approximate mental calculation is the capacity to estimate and manipulate numerical quantities, which is a phylogenetically ancient cognitive tool that humans share with other animals (Flombaum et al., 2005; Cantlon and Brannon, 2007; Piazza, 2010) and that arises early in life (Xu and Spelke, 2000; Izard et al., 2009). A widely accepted view (Dehaene, 1997) assumes that the mental representation of numerical magnitudes takes the form of an analog mental number line (hereafter, MNL). In the last decades, evidence has been collected to support the idea that on the MNL numerosities are spatially oriented in ascending order from left to right (Dehaene et al., 1993; Fias and Fischer, 2005; Hubbard et al., 2005; Rugani and Sartori, 2016; de Hevia et al., 2017). The SNARC effect (spatial numerical association of response codes; Dehaene et al., 1993) is often interpreted as evidence for the functional association between numbers and space: in a parity judgment tasks, where participant have to decide whether a displayed number is odd or even, left-hand responses are faster for relatively small number and right-hand responses for relatively large numbers (Dehaene et al., 1993; Fias and Fischer, 2005; Hubbard et al., 2005). Since the magnitude of the number is not relevant for the task, this spatial bias is assumed to reflect the automatic activation of the spatial mapping of magnitudes on the MNL (but for an alternative account see Santens and Gevers, 2008). The functional association between visuospatial processing and numerical magnitudes is additionally suggested by the mounting evidence showing that a shift of spatial attention

can be induced by number processing (Sallilas et al., 2008; Ranzini et al., 2015, 2016; for a review see Fischer and Knops, 2014). It is worth noting that a functional association also emerges between shifts of spatial attention and mental arithmetic (Masson and Pesenti, 2014, 2016; Mathieu et al., 2016, 2017; Masson et al., 2017a,b). Moreover, converging evidence from behavioral (Izard and Dehaene, 2008), computational (Dehaene and Changeux, 1993), and neurophysiological studies (Nieder and Miller, 2003) suggests that the MNL is logarithmically compressed, which means that the representational overlap between adjacent quantities increases proportionally to their size, in accordance with the Weber–Fechner law (see Piazza et al., 2010).

Approximate calculation also follows the Weber–Fechner law (Barth et al., 2006; Dehaene, 2007), but it also shows an additional response bias, that is the OM effect. Three mutually not exclusive mechanisms have been proposed to explain the OM effect: attentional shift account, heuristic account, and compression account. However, none of them aimed to describe how this effect changes over development. Evidence shows that the neural network that supports mental calculation undergoes substantial functional changes during development and reaches an adult-like configuration only during adolescence (Rosenberg-Lee et al., 2011; Soltanlou et al., 2017, 2018; Arsalidou et al., 2018; Peters and De Smedt, 2018). Therefore, in order to fully understand the cognitive mechanisms lying at the core of the OM effect it is important to measure its developmental dynamics and to evaluate whether the current accounts are able to explain these age-related changes. In what follows, we introduce these accounts of the OM effect and discuss the developmental trajectories predicted by each of them.

It has been proposed that mental calculation is grounded in neural circuits that originally evolved for processing visuospatial information (Anderson, 2007; Dehaene and Cohen, 2007; Knops et al., 2009a). Moreover, various evidence supports the existence of a functional relationship between visuospatial attention (i.e., shift of spatial attention) and mental calculation (Masson and Pesenti, 2014, 2016; Mathieu et al., 2016, 2017; Masson et al., 2017a,b). In line with these studies, the *attentional shift account* proposes that the OM effect is the result of this functional relationship (McCrink et al., 2007; Knops et al., 2009b; Pinheiro-Chagas et al., 2017). The central assumption of the attentional shift account hypothesizes that non-symbolic addition and subtraction are implemented by shifting spatial attention on a spatially oriented MNL. During approximate calculation, the first operand is mapped on the MNL, then the attentional focus shifts from the current position (i.e., the point corresponding to the magnitude of the first operand) to a new position (i.e., the point corresponding to the magnitude of the result) by a distance corresponding to the magnitude of the second operand. The OM effect is produced by a bias in the attentional shift, that is the attentional focus moves too far along the MNL in the direction of the operation, generating an overestimation and an underestimation of the result of addition and subtraction, respectively. Strong evidence for the hypothesis that visuospatial attention is co-opted during mental calculation is provided by the overlap in the posterior superior parietal lobule (PSPL)

of the neural activity associated with left/right saccades (i.e., visuospatial orientation) and mental calculation (Knops et al., 2009a).

McCrink and Wynn (2009) proposed the *heuristic account* to explain the finding that the OM effect also affects performance in 9 months old infants. This account assumes that infants adopted a simple heuristic to solve the problems: “if adding, accept larger outcomes,” “if subtracting, accept smaller outcomes.” For addition, this heuristic approach might encourage infants to perceive larger outcomes as more plausible compared smaller ones, and vice versa for subtraction. Recently, McCrink and Hubbard (2017) interpreted the finding that the OM effect increased in adults when available attentional resources were limited by dividing attention between two concurrent tasks as further evidence for the heuristics account. However, the heuristic account and the attentional shift account are deeply intertwined and can be considered as a single mechanism (i.e., heuristics-via-spatial-shifts account), that is the heuristic decision results from the visuospatial system (McCrink and Hubbard, 2017). Therefore, we will only focus on the attentional shift account, assuming that the two accounts provide equivalent predictions.

The attentional shift account has been developed to explain the OM effect in adults. Therefore, no predictions or hypotheses were proposed regarding how the attentional shifts on the MNL that accompany addition and subtraction emerge and whether they undergo substantial changes during development. Here, we propose that formal schooling (i.e., acquiring arithmetical skills) could reinforce (or even contribute to develop) the idea that addition is related with shifts toward larger numbers and subtraction toward smaller numbers. Namely, although mental calculation might be implemented as an attentional shift on the MNL before formal schooling, repeated exposition to spatial-numerical associations (e.g., the number line) might consolidate a systematic movement direction during the acquisition of arithmetical skills. Moreover, the systematic association between operations and results (i.e., when adding, the result is always larger than both operands; when subtracting, the result is always smaller than the first operand), that children are exposed to, could boost the attentional shift on the MNL. The influence of the attentional shift in the estimation of the result might increase with age and in turn a larger and more systematic bias would emerge. Therefore, one may predict an increasing OM effect during childhood. Moreover, it is worth noting that the co-opting of visuospatial attention during mental calculation seems to increase with age. In fact, significant functional changes associated with the neural activity elicited by symbolic arithmetic problem-solving have been found between 2nd and 3rd graders, that is 7–9 years old children (Rosenberg-Lee et al., 2011). During the processing of symbolic arithmetic problems, 3rd grade children showed greater activity in brain regions related to visuospatial attentional processes (posterior parietal cortex: intraparietal sulcus, superior parietal lobule, and angular gyrus) and high-order visual processing (ventral visual areas: lingual gyrus, right lateral occipital cortex, and right parahippocampal gyrus), compared to 2nd grade children.

The *compression account* has been proposed by McCrink et al. (2007) and deploys the logarithmic compression of the MNL to explain the OM effect. This compressed metric would generate a systematic operational bias in the direction of the operation due to the implementation of a linear arithmetic operation (i.e., addition or subtraction) on a logarithmically scaled mental representation. This mechanism acts in three steps. First, the operands are encoded as logarithmically compressed magnitudes on the MNL. Second, the logarithmic transformation is undone, which means that the operands are uncompressed to a linear scale. Third, the two uncompressed operands are added or subtracted. The OM effect results from the inaccuracy of the uncompression process. If the uncompression is ineffective the arithmetic operation is performed on logarithmic values and thus the generated outcome corresponds to an extreme overestimation or underestimation for addition and subtraction, respectively. If the uncompression is highly accurate the operation is performed on the linear scale, in which case the generated outcome corresponds (approximately) to the arithmetically correct result. A more plausible scenario is to assume that the actual degree of uncompression lies between these two extreme possibilities. An example can help describe this idea. If uncompression fails, adding two operands (e.g., 26 and 14) corresponds to adding their logarithmically compressed internal representation, that is $\log(26) \approx 3.26$ and $\log(14) \approx 2.64$, respectively. Since adding the logarithm of two numbers is equivalent to multiplying their linear values, the system generates an extreme overestimation of the correct result: $\log(26) + \log(14) \approx 5.9$, which in linear scale corresponds to $e^{5.9} \approx 26 \times 14 \approx 364$. However, the actual approximate addition performed by the system is much more accurate (see for example McCrink et al., 2007), and thus the uncompression is to some extent carried out and the generated outcome is much closer to the correct result. The same reasoning is valid to explain the mechanisms underpinning the underestimation of subtraction outcomes.

What developmental trajectory of the OM effect is expected according to the compression account? This account focuses on the logarithmic compression of the MNL. A large body of evidence suggests that the representational metric of the MNL shifts from a logarithmic to a linear scale during childhood (Siegler and Opfer, 2003; Siegler and Booth, 2004; Booth and Siegler, 2006, 2008; Laski and Siegler, 2007; Opfer and Siegler, 2007 but for a different interpretation see Barth and Paladino, 2011). The logarithmic-to-linear shift of the MNL implies that the compression of this magnitude representation decreases with age and probably with accumulation of experience in formal mathematics teaching. Therefore, the uncompression of the operands, performed before the approximate mental calculation, starts from a highly logarithmic scale in young children and from a more linear scale in adults. The degree of uncompression required to generate an accurate outcome is thus greater in young children and this in turn could lead to a stronger OM effect. The compression account therefore predicts that the size of the OM effect is higher in young children and decreases with age to reach an adult-like pattern in older children. It is worth noting that, as discussed below, the inverse OM effect (i.e., overestimation

of subtraction problems) found in 6/7 years old children (Knops et al., 2013) already provides evidence against this account.

MATERIALS AND METHODS

The sample and the tasks analyzed in the present paper were administered to children as part of a larger study conducted in Brazil (for a more precise description of this larger study see Pinheiro-Chagas et al., 2014).

Participants

One hundred seventy-two children from first to sixth grade were recruited from private and public schools in Brazil. Ten children were not able to perform non-symbolic numerical tasks, as shown by the fact that they failed to perform a non-symbolic number comparison task (this task is not reported here, for a more detailed description of this task see Pinheiro-Chagas et al., 2014). In that non-symbolic number comparison task, children had an accuracy less than 55% and a poor fit ($R^2 < 0.2$) in the estimation of the Weber fraction, and thus were excluded from the study. These ten children were also not included in the present analyses. The final sample consisted of 162 children (66 boys, 96 girls) between 8 and 12 years of age (mean = 9.7 years, $SD = 1.1$; 8 years old: 24 children, 9 years old: 54, 10 years old: 50, 11 years old: 20, 12 years old: 14). Informed written consent was obtained from the parents and oral consent from the children. This study was approved by the ethics review board of the Federal University of Minas Gerais, Brazil (COEP-UFGM).

All children performed above the 25th percentile in the spelling (mean = 110.08, $SD = 8.13$, range = [85, 126]) and arithmetic (mean = 108.92, $SD = 11.41$, range = [86, 134]) subtests of the TDE (Teste de Desempenho Escolar; Stein, 1994) and had a normal intelligence (mean = 110.61, $SD = 10.55$, range = [86, 134]), as measured by Raven's Colored Progressive Matrices (Angelini et al., 1999).

Tasks

Non-symbolic Estimation Task

In this task children were asked to estimate and report verbally the numerosity of a set of dots visually presented on a computer screen. Dots were displayed in black within a white circle, which was presented against a black background. The following numerosities were presented: 10, 16, 24, 32, 48, 56, or 64 dots. Each numerosity was presented five times (in a different configuration), resulting in a total of 35 trials. The same numerosity never appeared in consecutive trials. Each trial started with a fixation point (i.e., a white cross at the center of the screen) presented for 500 ms, followed by the onset of the set of dots which remained on the screen until spacebar was pressed or for up to 1000 ms. During the presentation of the dots, as soon as the child responded, the examiner, who was seated next to the child, pressed the spacebar on the keyboard and typed the child's answer. The next trial started after an intertrial interval of 700 ms, which consisted of a black screen. Dots were displayed on the screen for up to 1000 ms only to prevent counting. To prevent the use of non-numerical features, total dot area was

held constant across the trials and thus it could not be used as a clue to estimate the different numerosities. The average dot-size of the dots was selected so that the total area remained constant, but the dot-size of each dot could vary with a normal distribution with the mean selected to provide constant area across the trials. Therefore, while the average dot-size covaried negatively with numerosity, the dot-size of the single dots could not be used as a cue to evaluate the numerosity of the set. To avoid memorization effects due to the repetition of a specific numerosity, on each trial, the stimuli were randomly chosen from a set of 10 precomputed images with the given numerosity. To exclude extreme responses, the normalized mean estimated value was calculated for each child and each of the seven presented numerosities, then responses ± 3 SD from the mean estimated value were considered outliers and excluded from the analysis (3.5% of the trials). Children's number acuity was measured in term of individual mean coefficient of variation (i.e., separately for each numerosity, the ratio of standard deviation and mean chosen value).

Non-symbolic Approximate Calculation Task

This task has been adapted from Knops et al. (2013) study. Children were asked to solve approximate addition and subtraction problems with operands and proposed results presented in a non-symbolic notation (i.e., sets of dots). Problems are reported in **Table 1**. Eight addition and eight subtraction problems were generated. Both arithmetic operations had the same range of possible outcomes: 10, 16, 26, 40. To prevent the subjects from memorizing the problems, the operands were randomly "jittered" by adding a random value r , with $r \in J$ and $J = [-1, 0, 1]$. For each correct outcome, seven response alternatives were generated as round($c \times 2.5^{i/3}$), where c is the correct result and $i = [-3, -2, -1, 0, 1, 2, 3]$. To avoid a strategy of always selecting the response alternative falling in the middle of the proposed range, only five of the seven generated alternatives were presented in a trial (see **Table 1**). In one half of the trials, the presented responses were the upper five (henceforth, high range), and thus the correct outcome was the second smallest numerosity. In the other half, the presented responses were the lower five (henceforth, low range), and thus the correct outcome was the fourth smallest numerosity. Each trial was repeated twice and thus the total number of trials was 64: 2 operations (addition and subtraction) \times 8 problems \times 2 ranges (high and low) \times 2 repetitions. To prevent the use of non-numerical features, total dot area and dot-size were manipulated as in the non-symbolic estimation task. To avoid memorization effects due to the repetition of a specific numerosity, on each trial, the stimuli were randomly chosen from a set of 10 precomputed images with the given numerosity. Trials without response and trials where the selected response was ± 3 SD from the normalized mean chosen values (calculated combining addition and subtraction) were considered outliers and excluded from the analysis (3.1% of the trials). To analyze the OM effect, for each child and for each operation (addition vs. subtraction), mean chosen value, standard deviation, and coefficient of variation (i.e., the ratio of standard deviation and mean chosen value) were calculated for each of the four correct outcomes.

TABLE 1 | Operands, correct outcome (C) and deviant (D) outcomes presented in the non-symbolic arithmetic problems.

Operands		Correct results and deviant proposed outcomes						
		1/2.5	1/1.8	1/1.4	1	1.4	1.8	2.5
Addition								
5	5	4	5	7	10	14	18	25
6	4	4	5	7	10	14	18	25
8	8	6	9	12	16	22	29	40
10	6	6	9	12	16	22	29	40
13	13	10	14	19	26	35	48	65
18	8	10	14	19	26	35	48	65
20	20	16	22	29	40	54	74	100
26	14	16	22	29	40	54	74	100
Subtraction								
16	6	4	5	7	10	14	18	25
20	10	4	5	7	10	14	18	25
24	8	6	9	12	16	22	29	40
32	16	6	9	12	16	22	29	40
40	14	10	14	19	26	35	48	65
52	26	10	14	19	26	35	48	65
62	22	16	22	29	40	54	74	100
80	40	16	22	29	40	54	74	100
Range								
	Low	D	D	D	C	D		
	High			D	C	D	D	D

The last two rows report the set of outcomes presented in the two ranges.

To provide a child-friendly paradigm, problems were embedded in a story of a monkey having a box of balls (**Figure 1**). Each trial started with the drawing of the monkey's face presented for 500 ms. After the offset of the monkey's face, an empty brown box (against a black background) appeared at the bottom of the screen and a first set of red dots moved into the box. The first set of dots appeared at the top of the screen and moved toward the box until the dots disappeared inside it. For addition problems, a second set of red dots appeared at the top of the screen and disappeared inside the box in the same way. For subtraction problems, a set of red dots moved out of the box and disappeared at the top of the screen. Both for the first and the second sets, the duration of the dots movement (from the appearance to the disappearance) was 1000 ms. After the second set of dots disappeared, the box was replaced by the top-view of five boxes that contained five different sets of dots (i.e., five responses alternatives). Two boxes appeared on the left of the screen, two on the right, and one on the top. Children were asked to click with the left-key of the mouse on the box containing the set of dots which numerosity was the closest to the correct outcome of the operation. The beginning of the response active period was indicated by the appearance of the mouse pointer on top of a green star in the center of the screen. A training period consisting of two trials preceded the testing phase. In the training period, there was no time limit for the response and feedback was provided by a frame around the chosen box. The appearance of a green frame indicated a correct response, whereas a red frame indicated an incorrect response. If the response was incorrect, the child was asked to

choose another box, and this procedure was repeated until the correct box was chosen. Before testing phase, the children were asked if they had understood the task, and if not, the training was repeated until they confirmed that they understood the task. In the testing phase, children had a maximum of 10,000 ms to select the box and the chosen box was indicated by a neutral blue frame (i.e., no feedback provided). Addition and subtraction problems were presented in different blocks counterbalanced across participants.

Data Analysis

All analyses were performed using R-project software (R Core Team, 2015) and RStudio software (RStudio Team, 2015). In the following analyses, ANOVAs were Greenhouse-Geisser corrected (Greenhouse and Geisser, 1959) when the assumption of sphericity was violated; uncorrected degrees of freedom and epsilon values (ϵ_{GG}) are reported. In the *post hoc* analyses all *p*-values have been corrected with Holm's method (Holm, 1979). For the OM effect, effect sizes are reported following the recommendation of Lakens (2013). Additional analyses of children's performance (absolute error) and of the operational bias (ratio) are reported in the Appendix A.

RESULTS

The results of all the ANOVAs performed on the tasks are reported in the Appendix B (Supplementary Table S2).

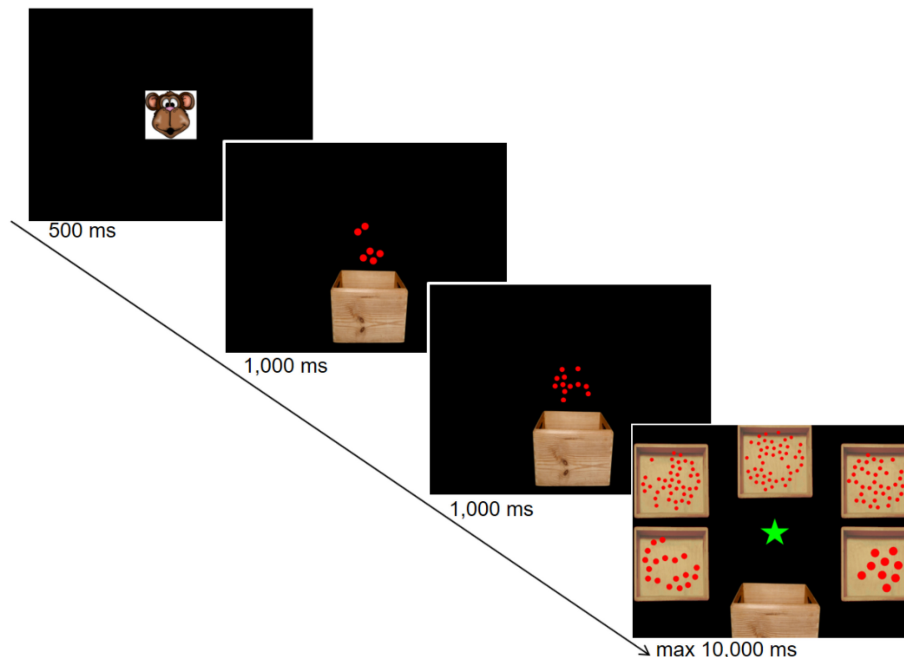


FIGURE 1 | Trial sequence of the non-symbolic approximate calculation task. The example shows the screenshots from a non-symbolic addition trial. During the response period, the five response alternatives were presented in a circle-like shape around the center of the screen (i.e., green star) with two boxes on the left of the screen, two on the right, and one on the top.

Non-symbolic Estimation Task

The first analysis aims to evaluate the performance of children in the non-symbolic number estimation task. Mean chosen numerosity and CV were analyzed with a repeated measure ANOVA with displayed numerosity (i.e., 10, 16, 24, 32, 48, 56, and 64 dots) as within-subject factor and age (i.e., 8 to 12 years old) as between-subject factor. Mean chosen numerosities significantly increased with displayed numerosity [$F(6,942) = 313.45$, $p < 0.001$, $\epsilon_{GG} = 0.27$, generalized $\eta^2 = 0.47$]. However, as shown in **Figure 2**, and in line with adults' behavior (Knops et al., 2014), children underestimated the larger displayed numerosities. To verify whether this pattern was statistically significant a repeated measure correlation (Bakdash and Marusich, 2017) was performed between numerical difference (chosen numerosity minus displayed numerosity) and displayed numerosity. There was a strong negative correlation between numerical difference and displayed numerosity [$r_{rm}(971) = -0.57$, 95% CI = $[-0.61, -0.53]$, $p < 0.001$], that is the discrepancy between displayed and chosen values increased with numerosity (**Figure 2**). In the ANOVA, neither the main effect of age nor the interaction was significant.

On the basis of the assumption that mental numerosity representation is subjected to the Weber-Fechner law, the CV should not covary with displayed numerosity (i.e., the CV should be constant across numerosities). As shown in **Figure 2**, the CV is lowest for the displayed numerosity 10 and increases with displayed numerosity [$F(6,942) = 11.04$, $p < 0.001$, $\epsilon_{GG} = 0.92$, $\eta_G^2 = 0.05$]. To further explore the relationship between CV and displayed numerosity, we performed a repeated measure

correlation (Bakdash and Marusich, 2017) between these two variables. A weak positive correlation emerged [$r_{rm}(971) = 0.16$, 95% CI = $[0.10, 0.22]$, $p < 0.001$], showing that the CV slightly increases with displayed numerosity. The ANOVA also revealed that the CV decreased with age [$F(4,157) = 5.26$, $p < 0.001$, $\eta_G^2 = 0.04$; see **Figure 2**] but no interaction was observed [$F(24, 942) < 1$]. This indicates that the overall accuracy increased with age.

To account for putative effects of inflated variance due to small number of trials in each displayed numerosity, we repeated these analyses using the z-transformed scores. For both mean chosen numerosity and CV, we calculated the standardized z-scores over all displayed numerosity for each child. The mean z-scores were entered into a repeated measure ANOVA with age as between-subject factor. Similar results emerged. In fact, age significantly influenced CV [$F(4,157) = 5.37$, $p < 0.001$] but not mean chosen numerosity [$F(4,157) < 1$].

Distribution of Responses in Approximate Addition and Subtraction

In each trial, the set of five proposed alternatives was sampled from either the lower range of responses (alternatives from 1 to 5, see **Table 1**) or the higher range (alternatives from 3 to 7, see **Table 1**). Therefore, the correct outcome was either the second (high range) and the fourth (low range) smaller proposed alternative. If children were able to solve the calculation, the response pattern should show a non-flat distribution centered on the correct outcome (i.e., second or fourth smaller alternative for high and low range, respectively).

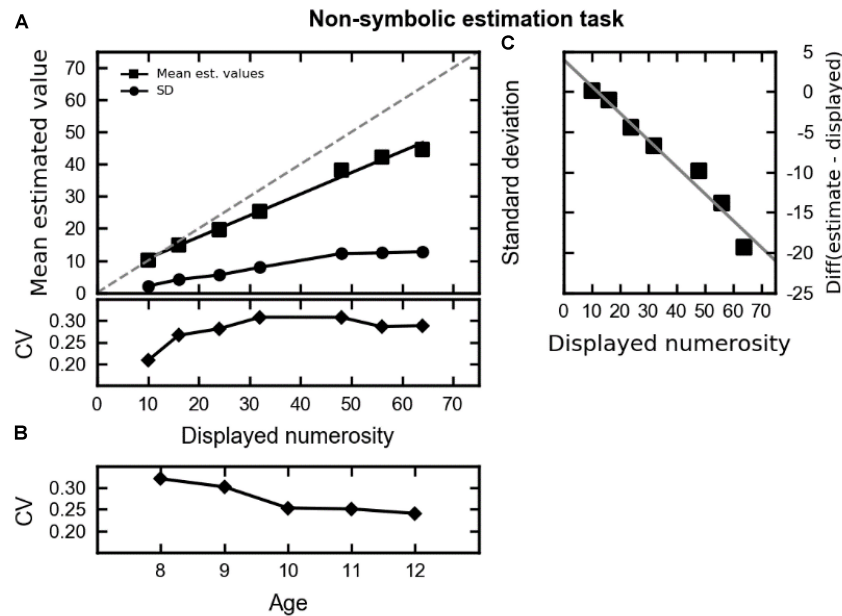


FIGURE 2 | (A) The top part shows the mean chosen numerosities (squares; the black line represents the regression model) and standard deviation (circles) plotted against the displayed numerosity. The gray dashed line represents perfect performance. The lower part reports the mean CV (coefficients of variation) plotted against the displayed numerosity. **(B)** The mean CV plotted against the age groups. **(C)** The difference between chosen numerosity and displayed numerosity plotted against the displayed numerosity. The gray line represents a regression model between the variables.

Mean (arcsine-transformed) percentage of choice was analyzed with a repeated-measure ANOVA with response category (i.e., 1 to 5), range (i.e., low vs. high), and operation (i.e., addition vs. subtraction) as within-subject factors and age (i.e., 8 to 12 years old) as between-subject factor. Results are reported in Supplementary Table S2 (see Appendix B). In particular, both the operation \times range \times response category interaction [$F(4,628) = 141.89$, $p < 0.001$, $\epsilon\text{GG} = 0.95$, generalized $\eta^2 = 0.16$] and the age \times range \times response category interaction [$F(16,628) = 1.71$, $p = 0.048$, $\epsilon\text{GG} = 0.89$, generalized $\eta^2 = 0.01$] were significant. Moreover, the four-way interaction showed a tendency toward significance [$F(16,628) = 1.54$, $p = 0.085$, $\epsilon\text{GG} = 0.95$, generalized $\eta^2 < 0.01$]. The tendency of the four-way interaction and **Figure 3** suggest that the performance was different in the two operations. Therefore, to further explore this pattern, two additional ANOVAs were performed on mean percentage of choice with response category and range as within-subject factors and age as between-subject factor, separately for addition and subtraction.

For addition, the main effect of response category was significant [$F(4,628) = 22.06$, $p < 0.001$, $\epsilon\text{GG} = 0.89$, generalized $\eta^2 = 0.06$]. Moreover, the age \times response category [$F(16,628) = 2.19$, $p = 0.007$, $\epsilon\text{GG} = 0.89$, generalized $\eta^2 = 0.03$], the range \times response category interaction [$F(4,628) = 223.06$, $p < 0.001$, $\epsilon\text{GG} = 0.87$, generalized $\eta^2 = 0.43$] and the three-way interaction [$F(16,628) = 2.07$, $p = 0.012$, $\epsilon\text{GG} = 0.87$, generalized $\eta^2 = 0.03$] were significant (**Figure 3**).

For subtraction, only the main effect of response category [$F(4,628) = 19.18$, $p < 0.001$, $\epsilon\text{GG} = 0.89$, generalized $\eta^2 = 0.07$] and the age \times response category interaction [$F(16,628) = 2.02$,

$p = 0.014$, $\epsilon\text{GG} = 0.89$, generalized $\eta^2 = 0.03$] were significant, whereas neither the range \times response category interaction [$F(4,628) = 2.07$, $p = 0.087$] nor the three-way interaction [$F(16,628) < 1$] reached significance (**Figure 3**). The response distribution for subtraction was flatter, showing that children found more difficult to perform approximate subtraction.

Children's Performance in Approximate Calculation

In order to evaluate children's performance in approximate addition and subtraction, mean chosen response and standard deviation were analyzed with a repeated-measure ANOVA with correct outcome (i.e., 10, 16, 26, and 40) and operation (i.e., addition vs. subtraction) as within-subject factors and age (i.e., 8–12 years old) as between-subject factor. For mean chosen response, the main effect of correct outcome was significant [$F(3,471) = 1685.80$, $p < 0.001$, $\epsilon\text{GG} = 0.60$, $\eta^2_{\text{G}} = 0.76$]. Mean chosen responses increased with correct outcome (mean responses: 12.0, 17.3, 24.1, and 32.9 for the outcomes 10, 16, 26, and 40, respectively). Mean chosen responses were greater for addition (mean = 23.2) than for subtraction (mean = 19.9) [$F(1,157) = 93.49$, $p < 0.001$, $\eta^2_{\text{G}} = 0.12$]. Moreover, all the two-way interactions were significant: correct outcome \times operation [$F(3,471) = 131.81$, $p < 0.001$, $\epsilon\text{GG} = 0.72$, $\eta^2_{\text{G}} = 0.12$], correct outcome \times age [$F(12,471) = 2.03$, $p = 0.049$, $\epsilon\text{GG} = 0.60$, $\eta^2_{\text{G}} = 0.01$], operation \times age [$F(4,157) = 6.24$, $p < 0.001$, $\eta^2_{\text{G}} = 0.04$]. Interestingly, the three-way interaction was also significant [$F(12,471) = 2.78$, $p = 0.004$, $\epsilon\text{GG} = 0.72$, $\eta^2_{\text{G}} = 0.01$]. As shown in **Figure 4**, mean chosen values were overestimated for addition compared to subtraction, and this difference was greater

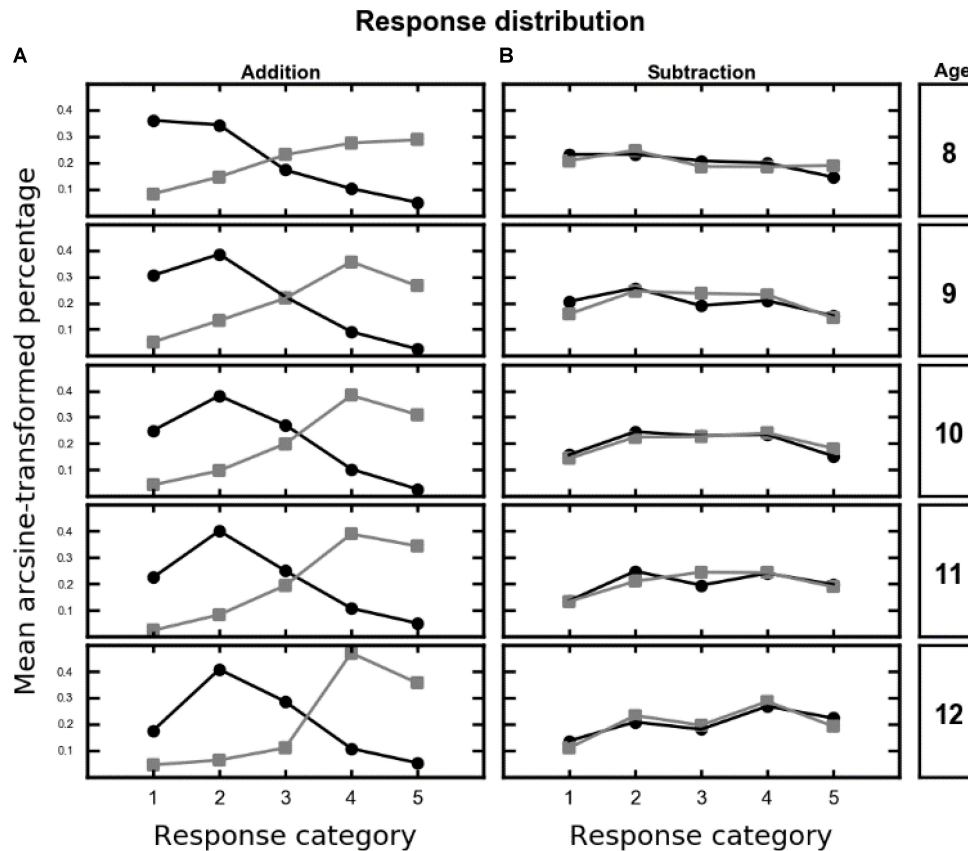


FIGURE 3 | Mean (arcsine-transformed) percentage of choice across the response category (x-axis) as a function of range (high: black circles, low: gray squares) and age (from 8 to 12, rows), for addition (A) and subtraction (B). For high range the correct outcome is the response category 2, for low range the correct outcome is the response category 4.

for larger numerosities and increased with age. This pattern reflects the OM effect and will be further investigated in the following section.

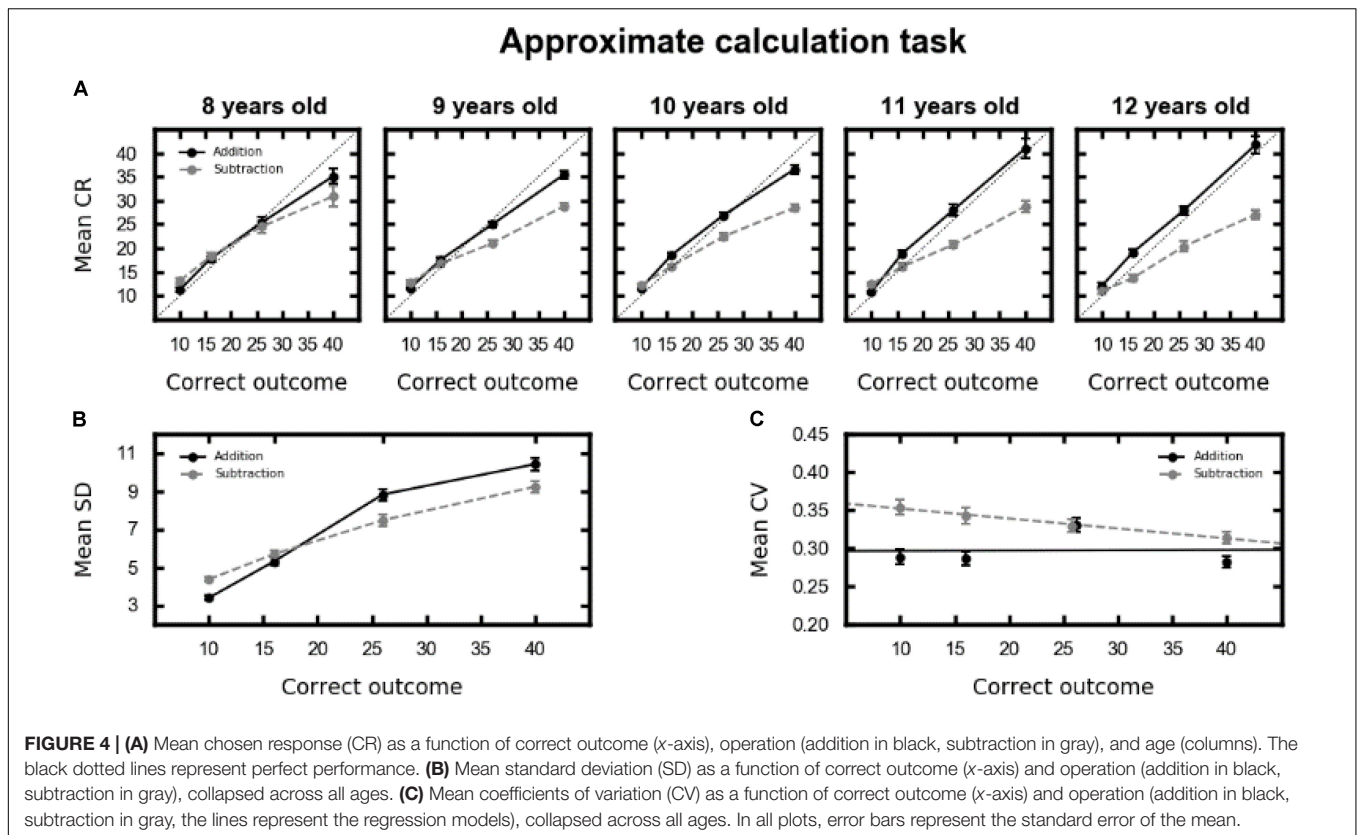
Standard deviation significantly increased with correct outcome [$F(3,471) = 275.66, p < 0.001, \epsilon_{GG} = 0.82, \eta^2_G = 0.35$]. However, this increase followed a different pattern in the two operations, as shown by the correct outcome by operation interaction [$F(3,471) = 18.17, p < 0.001, \epsilon_{GG} = 0.88, \eta^2_G = 0.02$], see **Figure 4**. No other main effects or interactions were significant.

To investigate whether children's mental numerosity representation follows Weber-Fechner law, a third ANOVA was performed on CV with correct outcome and operation as within-subject factors and age as between-subject factor. The main effect of correct outcome was significant [$F(3,471) = 5.88, p < 0.001, \epsilon_{GG} = 0.90, \eta^2_G = 0.01$] [outcomes 10: mean CV (SD) = 0.32 (0.09); outcome 16: 0.31 (0.09); outcome 26: 0.33 (0.09); outcome 40: 0.30 (0.07)]. Moreover, the CV was also significantly smaller for addition (mean = 0.30, SD = 0.08) than for subtraction (mean = 0.33, SD = 0.08) [$F(1,157) = 30.28, p < 0.001, \eta^2_G = 0.03$]. Finally, the interaction between correct outcome and operation was significant [$F(3,471) = 7.46, p < 0.001, \epsilon_{GG} = 0.96, \eta^2_G = 0.01$], see **Figure 4**. To further

investigate this interaction, we performed a repeated measure correlation between correct outcome and CV, separately for each operation. For addition, no correlation emerged between CV and correct outcome [$r_{rm}(485) = 0.005, 95\% \text{ CI} = [-0.08, 0.09], p = 0.91$]. For subtraction, a weak negative correlation emerged [$r_{rm}(485) = -0.17, 95\% \text{ CI} = [-0.25, -0.08], p < 0.001$], showing that mean CV slightly decreased with correct outcome, and thus the variability of the chosen response did not increase proportionally with the mean of the chosen response. These results are not perfectly consistent with the assumption that the underlying mental numerosity representation follows the Weber-Fechner law. However, since the CV did not covary with correct outcome in addition and only weakly correlated with it in subtraction (explained variance: 2.89%), the overall performance did not substantially deviate from this assumption.

Operational Momentum Effect

To investigate the developmental trajectory of the OM effect, the mean response bias was analyzed with a repeated-measure ANOVA with operation as within-subject factor and age as between-subject factor. Response bias was calculated as the mean difference between the logarithm of the chosen response and the logarithm of the correct outcome. Response bias was significantly

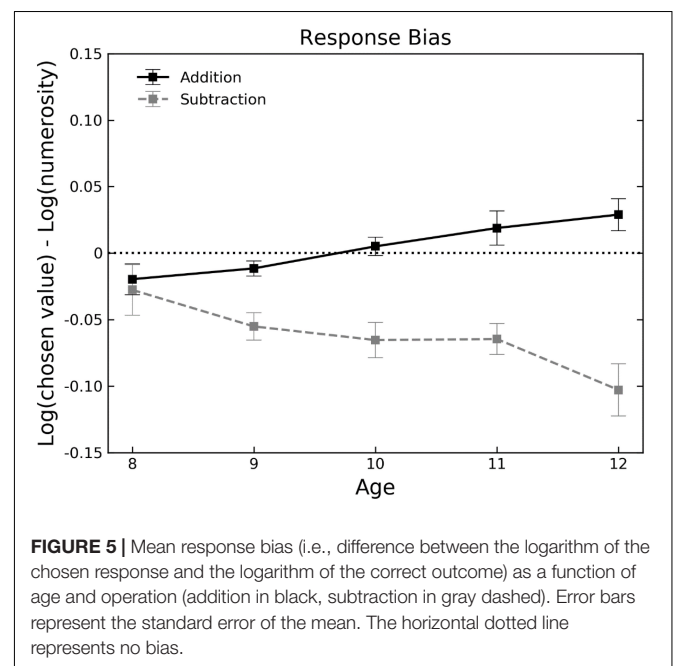


different between addition (-0.0004 , $SD = 0.05$) and subtraction (-0.06 , $SD = 0.08$) [$F(1,157) = 60.2$, $p < 0.001$, $\eta_G^2 = 0.17$]. The age by operation interaction was also significant [$F(4,157) = 4.45$, $p = 0.002$, $\eta_G^2 = 0.06$]. As shown in **Figure 5**, the OM effect monotonically increased with age¹, from no effect for younger children to a strong effect for older children (see **Table 2** for *post hoc* comparison and effect sizes). To further explore the addition and subtraction response biases separately, a second set of one-sample *t*-tests have been performed to evaluate whether they significantly differed from zero (biases significantly different from zero are shown in bold in **Table 2**). As shown in the table, only subtraction biases for the age groups from 9 to 12 were significantly different from zero [all *t*s < -4.97 , all *p*s < 0.01].

In Appendix A, we report an additional set of analyses that by and large confirms these findings.

DISCUSSION

This study aimed to investigate the developmental trajectory of the OM effect in children aged from 8 to 12 years old and to assess whether the current accounts are able to predict these age-related changes. Concerning the non-symbolic estimation



¹Since the sample size is unequal in the different age groups, we also performed two Spearman's correlation analyses between mean response bias and age (in months), separately for addition and subtraction. For addition, there was significant positive correlation [$r = 0.31$, $p < 0.001$]. For subtraction, there was significant negative correlation [$r = -0.24$, $p = 0.002$].

task, consistent with previous research (Izard and Dehaene, 2008; Knops et al., 2014; but for overestimation see Mejias and Schiltz, 2013), children underestimated the cardinality of displayed numerosities and this underestimation increased

TABLE 2 | *T*-tests comparing the response bias between addition and subtraction in the different age groups.

Age group	N	Addition		Subtraction		<i>t</i>	df	<i>p</i> -value	Cohen's <i>d_z</i>	Hedges' <i>g_{av}</i>
		Mean	SD	Mean	SD					
8	24	−0.020	0.057	−0.028	0.094	0.4	23	>0.1	0.08	0.10
9	54	−0.012	0.041	−0.055	0.075	3.61	53	0.005	0.49	0.71
10	50	0.005	0.048	−0.065	0.093	4.55	49	<0.001	0.64	0.94
11	20	0.019	0.058	−0.065	0.052	4.52	19	0.002	1.01	1.46
12	14	0.029	0.045	−0.103	0.073	5.04	13	0.002	1.35	2.04

All *p*-values have been corrected with Holm's method. For the calculation of the effect sizes (Cohen's *d_z* and Hedges' *g_{av}*) refers to Lakens (2013). Mean response biases significantly different from zero (i.e., one-sample *t*-tests, separately computed for each operation and age group) are in bold, all *ps* < 0.01.

with numerosity. Although the CV significantly increased with numerosity, the correlation between the two variable was weak ($r_{\text{rm}} = 0.16$). Moreover, both mean estimated values and standard deviation increased with displayed numerosity. This suggests that children's performance was by and large well captured by Weber–Fechner law, even if the CV was not perfectly linear across the entire numerical range. In line with previous findings that suggest that the Weber fraction decreases with age (Piazza et al., 2010; Halberda et al., 2012), the coefficient of variation also significantly decreased with age. Deviations may be due to non-numerical features of the stimulus set, for example. Further studies are needed to fully explain these inconsistencies.

In the approximate addition task, the distribution of responses clearly peaked around the correct outcome showing that children were able to solve these problems. The response distribution for subtraction problems, however, showed a different pattern. The distribution was flat for younger children (8 years old, see **Figure 3**) and in general the two ranges (low vs. high, see **Table 1**) were almost overlapped. Therefore, children found subtraction problems more difficult to solve compared to addition problems, in line with adults (Knops et al., 2009b). However, for subtraction problems, the significant main effect of response category and **Figure 3** suggest that children (at least in the age groups from 9 to 12) did not respond at random but rather selected more often values in the center of the response category range (i.e., 2, 3, 4) compared to the extremes (i.e., 1 and 5). This suggests that children might have used a different strategy to perform subtraction compared to addition. Despite the lower performance on subtractions problems, a clear OM effect emerged in our sample. Importantly, for addition the increase of the OM effect was accompanied by an increase in overall accuracy (see **Figure 3**). That is, while younger children made more and non-systematic errors, older children made less but systematic errors. Interestingly, the OM effect monotonically increases with age. While no effect was present in younger children (8 years-olds), the OM effect (i.e., the relative difference between the estimated responses in addition and subtraction) increased with age. In what follows, we first summarize the findings related to the evolution of the OM effect during childhood, and then we will discuss the implications of these findings for the current accounts of the OM effect (i.e., compression account and attentional shift account).

McCrink and Wynn (2009) found that 9 months old infants exhibit an OM effect similar to that found in adults. Although

the similarity between the OM effect found in infants (McCrink and Wynn, 2009) and adults (McCrink et al., 2007; Knops et al., 2009b) would suggest that the OM effect results from inherited mechanisms (since infants are not yet affected by cultural practices) and remains constant during development, a more complex pattern emerges if we consider a previous study (Knops et al., 2013) and the findings reported in the current paper. In fact, contrary to the expected continuity of the OM effect during development, Knops et al. (2013) found an inverse OM effect in 6/7 years old children: subtraction was significantly overestimated compared to addition. Finally, our results showed a monotonic increase of the OM effect with age. This complex developmental pattern indicates that the evolution of the OM effect is not linear. In fact, a standard OM effect emerges in infants (McCrink and Wynn, 2009), an inverse OM effect was found in 6/7 years old children (Knops et al., 2013), and our results show no OM in 8 years old children and a monotonically increasing OM effect from 9 to 12 years old.

How well do the current accounts predict the developmental-related changes of the OM effect? The *compression account* (McCrink et al., 2007) predicts that, due the logarithmic-to-linear shift of the MNL during childhood (Siegler and Opfer, 2003; Siegler and Booth, 2004; Booth and Siegler, 2006, 2008; Laski and Siegler, 2007; Opfer and Siegler, 2007; but for a different perspective see Barth and Paladino, 2011), the OM effect decreases with age. Our result clearly points in the opposite direction showing an increase of the OM effect.

In line with the recycling theory (Dehaene and Cohen, 2007; see also the redeployment theory, Anderson, 2007), which proposes that arithmetic calculation is grounded on the recycling of neural circuits that originally evolved for processing visuospatial information, the *attentional shift account* assumes that the OM effect is driven by the functional relationship between visuospatial attention and mental arithmetic. Strong evidence for the idea that visuospatial attention is co-opted during mental calculation is provided by the fact that the neural activity associated with left/right saccades (i.e., visuospatial orientation) and mental calculation overlap in the posterior superior parietal lobule (Knops et al., 2009a). Using fMRI data, these authors showed that a multivariate classifier algorithm trained to classify the neural activity elicited by leftward and rightward saccades was able to generalize to approximate arithmetic. Without further training, this algorithm was able to distinguish between addition and subtraction

by classifying approximate additions as rightward saccades. The activation of the same neural areas during rightward saccades and approximate addition speaks in favor of the recruitment of attentional shift mechanisms during mental calculation. This hypothesis stipulates a functional coupling between eye movements and arithmetic. A recent study provided confirmatory evidence for this notion (Klein et al., 2014). Participants' eye movements after the first saccade were observed to move to the right during addition problems and to the left in subtraction problems when asked to indicate the location of the result on a labeled line (Klein et al., 2014). Moreover, the redeployment of visuospatial attention during mental calculation seems to be enhanced during formal schooling (Rosenberg-Lee et al., 2011). Finally, on the behavioral level, too, even if spatial-numerical association already emerges in preschoolers, the evidence is mixed. For example, White et al. (2012) found that the SNARC effect emerged during the 2nd year of schooling in British students, that is at around 7 years of age, while 6-year-olds did not show a significant SNARC effect (see also Gibson and Maurer, 2016). Moreover, Yang et al. (2014) found a SNARC effect in kindergarteners (age range: 4.8–6.4 years), 2nd, 3rd, 5th, and 6th graders, while 1st and 4th graders did not show a significant effect (see also Patro and Haman, 2012). Hoffmann et al. (2013) also found mixed evidence for the emergence of the SNARC effect. While all children in the second-term (mean age: 5.8 years old) showed a SNARC effect, in the first-term group (5.5 years old) the effect emerged when a magnitude comparison task preceded a digit color judgment task but not when the task order was inverted. Moreover, in the magnitude comparison task the size of the SNARC effect was related to proficiency with Arabic numbers. This developmental pattern suggests that the spatial-numerical association is still immature in young children. We propose that formal schooling could bolster spatial-numerical associations and hence reinforce movement direction during addition (toward larger numbers) and subtraction (toward smaller numbers). Attentional shifts may implement the core cognitive function to carry out the shifts along the spatial mental number representation and may be affected in at least two ways by the emerging spatial-numerical associations. Either the amount of displacement in the direction of the operation on the MNL increases (i.e., generate a larger and/or more systematic bias) or the variance of displacement is reduced while the overall amplitude remains constant. Therefore, the attentional shift account predicts an increasing OM effect during childhood. Consistent with this prediction, we found a monotonous increase of the OM effect with age.

Although the attentional shift account is consistent with our results, a more complex picture emerges if the results from previous studies are taken into account. In fact, the inverse OM effect found in 6/7 years old children (Knops et al., 2013) is neither explained nor predicted by this account. However, Knops et al. (2013) showed that the direction of the OM effect was related to reorienting attention in a Posner paradigm. The reorientation effect was calculated as the difference in reaction times between valid (i.e., the target stimulus appeared on the left or right of a bidirectional arrow previously presented in the center of the screen) and invalid trials (i.e., the target

stimulus appeared opposite the pointing direction of a single-headed arrow). In their study, children who exhibited a smaller reorientation effect (i.e., more proficient to reorient attention after an invalid cue) also had a more regular OM effect (i.e., addition overestimated compare to subtraction). As those authors suggested, it can be hypothesized that the OM effect relies on a fully developed attentional system and on a robust functional association between visuospatial attention and mental calculation. Alternatively, it may suggest that inhibitory control of saccadic eye movements plays a crucial role for the association between attention and arithmetic. We can only speculate as to why an inverse OM effect emerges in 6/7 years old children and the youngest age group of our sample does not show any effect. The more immature attentional system (Rueda et al., 2004; Konrad et al., 2005) and the weaker functional connection between visuospatial processing and mental calculation (Rosenberg-Lee et al., 2011) might be at the origin of the inverse OM effect and its absence in younger children. Namely, the implementation of approximate addition and subtraction would not be yet supported by operation-specific, systematic attentional shifts on the MNL that produce misestimation in the direction of the operation.

The presence of a standard OM effect in infants (McCrink and Wynn, 2009) challenges the idea that the OM effect monotonically increases during childhood due to the consolidation of the engagement of visuospatial processing during mental calculation. However, this contradiction strongly relies on the idea that the development of cognitive performance always reflects linear developmental trajectories. However, as put forward by Siegler (1996), behavior may reflect the prevalence of heuristics and biases that wax and wane over time. That is, while infants may respond according to a given heuristic, the very same heuristic may be less influential during later periods in life. In children, performance in approximate calculation tasks may be performed with the support of the visuospatial system (i.e., the shift of the attentional focus on the MNL), while in infants the heuristic decision may result from simpler processes rather than from more sophisticated attentional mechanisms. Namely, in children (or adults) and infants the heuristic decision might result from different mechanisms. However, more evidence on the development of the OM effect is needed to unravel the cognitive mechanisms that drive the OM at different ages.

This study has some limitations. First, children's performance in subtraction was low compared to addition. The higher difficulty to estimate the result of approximate subtraction could be due to the use of different strategies to perform the two operations. To better understand how children perform approximate calculation, future research should further investigate this difference in performance. Second, despite the fairly large sample, 6/7 years old children were not included, that is the age group that showed the inverse OM effect. Future studies should include a larger age range in order to confirm the inverse OM effect and to further investigate the development of this effect. Third, we did not include any task to measure visuospatial attention. Future studies should investigate whether there is a correlation between the developmental trajectories of visuospatial attention and of the OM effect. Finally, the effect of education

is also accompanied by the maturation of neural network that supports mental calculation. In the analysis we focused on age, future research, however, should also disentangle the influence of age (neural maturation) and grade (education) on the OM effect. These two independent factors could make distinct contribution at various stages of development.

To sum up, we provided a novel finding on the developmental trajectory of the OM effect in children from 8 to 12 years old. The OM effect monotonically increases with age. This developmental pattern is inconsistent with the compression account. On the other hand, the attentional shift account provides a possible explanation of these results based on the functional relationship between visuospatial attention and mental calculation and on the effect of the acquisition of arithmetical skills during formal schooling. The attentional shift account leads to new predictions about a correlation between visuospatial processing and mental calculation which can be addressed in future studies. Our results provide an important empirical constraint to further explore the origin of the OM effect.

ETHICS STATEMENT

This study was carried out in accordance with the recommendations of ethics review board of the Federal University of Minas Gerais, Brazil (COEP-UFMG) with written informed consent from all subjects. All subjects gave written

informed consent in accordance with the Declaration of Helsinki. Informed written consent was obtained from the parents and oral consent from the children. The protocol was approved by the ethics review board of the Federal University of Minas Gerais, Brazil (COEP-UFMG).

AUTHOR CONTRIBUTIONS

PP-C, VH, GW, and AK designed the research. PP-C performed the research. DD and PP-C analyzed the data. DD drafted the manuscript. DD, PP-C, AK, VH, and GW contributed to write and revise the paper.

ACKNOWLEDGMENTS

We acknowledge support by the German Research Foundation (DFG) and the Open Access Publication Fund of Humboldt-Universität zu Berlin.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.01062/full#supplementary-material>

REFERENCES

- Anderson, M. L. (2007). Evolution of cognitive function via redeployment of brain areas. *Neuroscientist* 13, 13–21. doi: 10.1177/1073858406294706
- Angelini, A. L., Alves, I. C. B., Custódio, E. M., Duarte, W. F., and Duarte, J. L. M. (1999). *Matrizes Progressivas Coloridas de Raven – Escala Especial*. São Paulo: Centro Editor de Testes e Pesquisas em Psicologia, 435.
- Arsalidou, M., Pawliw-Levac, M., Sadeghi, M., and Pascual-Leone, J. (2018). Brain areas associated with numbers and calculations in children: meta-analyses of fMRI studies. *Dev. Cogn. Neurosci.* 30, 239–250. doi: 10.1016/j.dcn.2017.08.002
- Bakdash, J. Z., and Marusich, L. R. (2017). Repeated measures correlation. *Front. Psychol.* 8:456. doi: 10.3389/fpsyg.2017.00456
- Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., and Spelke, E. (2006). Non-symbolic arithmetic in adults and young children. *Cognition* 98, 199–222. doi: 10.1016/j.cognition.2004.09.011
- Barth, H. C., and Paladino, A. M. (2011). The development of numerical estimation: evidence against a representational shift. *Dev. Sci.* 14, 125–135. doi: 10.1111/j.1467-7687.2010.00962.x
- Booth, J. L., and Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Dev. Psychol.* 41, 189–201. doi: 10.1037/0012-1649.41.6.189
- Booth, J. L., and Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Dev.* 79, 1016–1031. doi: 10.1111/j.1467-8624.2008.01173.x
- Brown, S. W. (1985). Time perception and attention: the effects of prospective versus retrospective paradigms and task demands on perceived duration. *Percept. Psychophys.* 38, 115–124. doi: 10.3758/BF03198848
- Brown, S. W. (1997). Attentional resources in timing: Interference effects in concurrent temporal and nontemporal working memory tasks. *Percept. Psychophys.* 59, 1118–1140. doi: 10.3758/BF03205526
- Cantlon, J. F., and Brannon, E. M. (2007). Basic math in monkeys and college students. *PLoS Biol.* 5:e328. doi: 10.1371/journal.pbio.0050328
- Chen, Q., and Verguts, T. (2012). Spatial intuition in elementary arithmetic: a neurocomputational account. *PLoS One* 7:e31180. doi: 10.1371/journal.pone.0031180
- de Hevia, M. D., Veggioni, L., Streri, A., and Bonn, C. D. (2017). At birth, humans associate “Few” with Left and “Many” with right. *Curr. Biol.* 27, 3879–3884. doi: 10.1016/j.cub.2017.11.024
- Dehaene, S. (1997). *The Number Sense*. New York, NY: Oxford University Press.
- Dehaene, S. (2007). “Symbols and quantities in parietal cortex: Elements of a mathematical theory of number representation and manipulation,” in *Sensorimotor Foundations of Higher Cognition: Attention and Performance*, eds P. Haggard, Y. Rossetti, and W. Kawato (Oxford: Oxford University Press), 527–574.
- Dehaene, S., Bossini, S., and Giraux, P. (1993). The mental representation of parity and number magnitude. *J. Exp. Psychol. Gen.* 122, 371–396. doi: 10.1037/0096-3445.122.3.371
- Dehaene, S., and Changeux, J. P. (1993). Development of elementary numerical abilities: a neuronal model. *J. Cogn. Neurosci.* 5, 390–407. doi: 10.1162/jocn.1993.5.4.390
- Dehaene, S., and Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron* 56, 384–398. doi: 10.1016/j.neuron.2007.10.004
- Fias, W., and Fischer, M. H. (2005). “Spatial representation of number,” in *Handbook of Mathematical Cognition*, ed. J. I. D. Campbell (Hove: Psychology Press), 43–54.
- Fischer, M. H., and Knops, A. (2014). Attentional cueing in numerical cognition. *Front. Psychol.* 5:1381. doi: 10.3389/fpsyg.2014.01381
- Flombaum, J. I., Junge, J. A., and Hauser, M. D. (2005). Rhesus monkeys (*Macaca mulatta*) spontaneously compute addition operations over large numbers. *Cognition* 97, 315–325. doi: 10.1016/j.cognition.2004.09.004
- Gibson, L. C., and Maurer, D. (2016). Development of SNARC and distance effects and their relation to mathematical and visuospatial abilities. *J. Exp. Child Psychol.* 150, 301–313. doi: 10.1016/j.jecp.2016.05.009
- Glicksohn, J., and Hadad, Y. (2012). Sex differences in time production revisited. *J. Individ. Differ.* 33, 35–42. doi: 10.1027/1614-0001/a000059

- Greenhouse, S. W., and Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika* 24, 95–112. doi: 10.1007/BF02289823
- Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., and Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proc. Natl. Acad. Sci. U.S.A.* 109, 11116–11120. doi: 10.1073/pnas.1200196109
- Hoffmann, D., Hornung, C., Martin, R., and Schiltz, C. (2013). Developing number–space associations: SNARC effects using a color discrimination task in 5-year-olds. *J. Exp. Child Psychol.* 116, 775–791. doi: 10.1016/j.jecp.2013.07.013
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 6, 65–70.
- Hubbard, E. M., Piazza, M., Pinel, P., and Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nat. Rev. Neurosci.* 6, 435–448. doi: 10.1038/nrn1684
- Izard, V., and Dehaene, S. (2008). Calibrating the mental number line. *Cognition* 106, 1221–1247. doi: 10.1016/j.cognition.2007.06.004
- Izard, V., Sann, C., Spelke, E. S., and Streri, A. (2009). Newborn infants perceive abstract numbers. *Proc. Natl. Acad. Sci. U.S.A.* 106, 10382–10385. doi: 10.1073/pnas.0812142106
- Klein, E., Huber, S., Nuerk, H. C., and Moeller, K. (2014). Operational momentum affects eye fixation behaviour. *Q. J. Exp. Psychol.* 67, 1614–1625. doi: 10.1080/17470218.2014.902976
- Knops, A., Dehaene, S., Berteletti, I., and Zorzi, M. (2014). Can approximate mental calculation account for operational momentum in addition and subtraction? *Q. J. Exp. Psychol.* 67, 1541–1556. doi: 10.1080/17470218.2014.890234
- Knops, A., Thirion, B., Hubbard, E. M., Michel, V., and Dehaene, S. (2009a). Recruitment of an area involved in eye movements during mental arithmetic. *Science* 324, 1583–1585. doi: 10.1126/science.1171599
- Knops, A., Viarouge, A., and Dehaene, S. (2009b). Dynamic representations underlying symbolic and nonsymbolic calculation: evidence from the operational momentum effect. *Atten. Percept. Psychophys.* 71, 803–821. doi: 10.3758/APP.71.4.803
- Knops, A., Zitzmann, S., and McCrink, K. (2013). Examining the presence and determinants of operational momentum in childhood. *Front. Psychol.* 4:325. doi: 10.3389/fpsyg.2013.00325
- Konrad, K., Neufang, S., Thiel, C. M., Specht, K., Hanisch, C., Fan, J., et al. (2005). Development of attentional networks: an fMRI study with children and adults. *Neuroimage* 28, 429–439. doi: 10.1016/j.neuroimage.2005.06.065
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front. Psychol.* 4:863. doi: 10.3389/fpsyg.2013.00863
- Laski, E., and Siegler, R. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child Dev.* 78, 1723–1743. doi: 10.1111/j.1467-8624.2007.01087.x
- Lindemann, O., and Tira, M. (2011). Operational momentum in numerosity production judgments of multi-digit number problems. *J. Psychol.* 219, 50–57. doi: 10.1027/2151-2604/a000046
- Marghetis, T., Núñez, R., and Bergen, B. K. (2014). Doing arithmetic by hand: hand movements during exact arithmetic reveal systematic, dynamic spatial processing. *Q. J. Exp. Psychol.* 67, 1579–1596. doi: 10.1080/17470218.2014.897359
- Masson, N., and Pesenti, M. (2014). Attentional bias induced by solving simple and complex addition and subtraction problems. *Q. J. Exp. Psychol.* 67, 1514–1526. doi: 10.1080/17470218.2014.903985
- Masson, N., and Pesenti, M. (2016). Interference of lateralized distractors on arithmetic problem solving: a functional role for attention shifts in mental calculation. *Psychol. Res.* 80, 640–651. doi: 10.1007/s00426-015-0668-7
- Masson, N., Pesenti, M., Coyette, F., Andres, M., and Dormal, V. (2017a). Shifts of spatial attention underlie numerical comparison and mental arithmetic: evidence from a patient with right unilateral neglect. *Neuropsychology* 31, 822–833. doi: 10.1037/neu0000361
- Masson, N., Pesenti, M., and Dormal, V. (2017b). Impact of optokinetic stimulation on mental arithmetic. *Psychol. Res.* 81, 840–849. doi: 10.1007/s00426-016-0784-z
- Mathieu, R., Epinat-Duclos, J., Léone, J., Fayol, M., Thevenot, C., and Prado, J. (2017). Hippocampal spatial mechanisms relate to the development of arithmetic symbol processing in children. *Dev. Cogn. Neurosci.* 30, 324–332. doi: 10.1016/j.dcn.2017.06.001
- Mathieu, R., Gourjon, A., Couderc, A., Thevenot, C., and Prado, J. (2016). Running the number line: rapid shifts of attention in single-digit arithmetic. *Cognition* 146, 229–239. doi: 10.1016/j.cognition.2015.10.002
- McCrink, K., Dehaene, S., and Dehaene-Lambertz, G. (2007). Moving along the number line: operational momentum in nonsymbolic arithmetic. *Percept. Psychophys.* 69, 1324–1333. doi: 10.3758/BF03192949
- McCrink, K., and Hubbard, T. (2017). Dividing attention increases operational momentum. *J. Numer. Cogn.* 3, 230–245. doi: 10.5964/jnc.v3i2.34
- McCrink, K., and Wynn, K. (2009). Operational momentum in large number addition and subtraction by 9-month-olds. *J. Exp. Child Psychol.* 103, 400–408. doi: 10.1016/j.jecp.2009.01.013
- Mejias, S., and Schiltz, C. (2013). Estimation abilities of large numerosities in Kindergartners. *Front. Psychol.* 4:518. doi: 10.3389/fpsyg.2013.00518
- Mioni, G., Stablum, F., McClintock, S. M., and Grondin, S. (2014). Different methods for reproducing time, different results. *Atten. Percept. Psychophys.* 76, 675–681. doi: 10.3758/s13414-014-0625-3
- Nieder, A., and Miller, E. K. (2003). Coding of cognitive magnitude: compressed scaling of numerical information in the primate prefrontal cortex. *Neuron* 37, 149–157. doi: 10.1016/S0896-6273(02)01144-3
- Opfer, J., and Siegler, S. (2007). Representational change and children's numerical estimation. *Cogn. Psychol.* 55, 169–195. doi: 10.1016/j.cogpsych.2006.09.002
- Patro, K., and Haman, M. (2012). The spatial–numerical congruity effect in preschoolers. *J. Exp. Child Psychol.* 111, 534–542. doi: 10.1016/j.jecp.2011.09.006
- Peters, L., and De Smedt, B. (2018). Arithmetic in the developing brain: a review of brain imaging studies. *Dev. Cogn. Neurosci.* 30, 265–279. doi: 10.1016/j.dcn.2017.05.002
- Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends Cogn. Sci.* 14, 542–551. doi: 10.1016/j.tics.2010.09.008
- Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., et al. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition* 116, 33–41. doi: 10.1016/j.cognition.2010.03.012
- Pinhas, M., and Fischer, M. H. (2008). Mental movements without magnitude? A study of spatial biases in symbolic arithmetic. *Cognition* 109, 408–415. doi: 10.1016/j.cognition.2008.09.003
- Pinheiro-Chagas, P., Dotan, D., Piazza, M., and Dehaene, S. (2017). Finger tracking reveals the covert stages of mental arithmetic. *Open Mind* 1, 30–41. doi: 10.1162/OPMI_a_00003
- Pinheiro-Chagas, P., Wood, G., Knops, A., Krinzinger, H., Lonnemann, J., Starling-Alves, I., et al. (2014). In how many ways is the approximate number system associated with exact calculation? *PLoS One* 9:e111155. doi: 10.1371/journal.pone.0111155
- R Core Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Ranzini, M., Lisi, M., Blini, E., Pitteri, M., Treccani, B., Priftis, K., et al. (2015). Larger, smaller, odd or even? Task-specific effects of optokinetic stimulation on the mental number space. *J. Cogn. Psychol.* 27, 459–470. doi: 10.1080/20445911.2014.941847
- Ranzini, M., Lisi, M., and Zorzi, M. (2016). Voluntary eye movements direct attention on the mental number space. *Psychol. Res.* 80, 389–398. doi: 10.1007/s00426-015-0741-2
- Rosenberg-Lee, M., Barth, M., and Menon, V. (2011). What difference does a year of schooling make? Maturation of brain response and connectivity between 2nd and 3rd grades during arithmetic problem solving. *Neuroimage* 57, 796–808. doi: 10.1016/j.neuroimage.2011.05.013
- RStudio Team (2015). *RStudio: Integrated Development for R*. Boston, MA: RStudio, Inc.
- Rueda, M. R., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., et al. (2004). Development of attentional networks in childhood. *Neuropsychologia* 42, 1029–1040. doi: 10.1016/j.neuropsychologia.2003.12.012
- Rugani, R., and Sartori, L. (2016). Numbers in action. *Front. Hum. Neurosci.* 10:388. doi: 10.3389/fnhum.2016.00388

- Sallilas, E., El Yagoubi, R., and Semenza, C. (2008). Sensory and cognitive processes of shifts of spatial attention induced by numbers: an ERP study. *Cortex* 44, 406–413. doi: 10.1016/j.cortex.2007.08.006
- Santens, S., and Gevers, W. (2008). The SNARC effect does not imply a mental number line. *Cognition* 108, 263–270. doi: 10.1016/j.cognition.2008.01.002
- Siegler, R. S. (1996). *Emerging Minds: The Process of Change in Children's Thinking*. Oxford: Oxford University Press.
- Siegler, R. S., and Booth, J. (2004). Development of numerical estimation in young children. *Child Dev.* 75, 428–444. doi: 10.1111/j.1467-8624.2004.00684.x
- Siegler, R. S., and Opfer, J. (2003). The development of numerical estimation: evidence for multiple representations of numerical quantity. *Psychol. Sci.* 14, 237–243. doi: 10.1111/1467-9280.02438
- Soltanlou, M., Artemenko, C., Dresler, T., Haeussinger, F. B., Fallgatter, A. J., Ehlis, A. C., et al. (2017). Increased arithmetic complexity is associated with domain-general but not domain-specific magnitude processing in children: a simultaneous fNIRS-EEG study. *Cogn. Affect. Behav. Neurosci.* 17, 724–736. doi: 10.3758/s13415-017-0508-x
- Soltanlou, M., Artemenko, C., Ehlis, A. C., Huber, S., Fallgatter, A. J., Dresler, T., et al. (2018). Reduction but no shift in brain activation after arithmetic learning in children: a simultaneous fNIRS-EEG study. *Sci. Rep.* 8:1707. doi: 10.1038/s41598-018-20007-x
- Stein, L. M. (1994). *TDE – Teste do Desempenho Escolar. Manual Para Aplicação e Interpretação*. São Paulo: Casa do Psicólogo, 204.
- White, S., Szucs, D., and Soltész, F. (2012). Symbolic number: The integration of magnitude and spatial representations in children aged 6 to 8 years. *Front. Psychol.* 2:392. doi: 10.3389/fpsyg.2011.00392
- Wynn, K. (1992). Addition and subtraction by human infants. *Nature* 358, 749–750. doi: 10.1038/358749a0
- Xu, F., and Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition* 74, B1–B11. doi: 10.1016/S0010-0277(99)00066-9
- Yang, T., Chen, C., Zhou, X., Xu, J., Dong, Q., and Chen, C. (2014). Development of spatial representation of numbers: a study of the SNARC effect in Chinese children. *J. Exp. Child Psychol.* 117, 1–11. doi: 10.1016/j.jecp.2013.07.013

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Pinheiro-Chagas, Didino, Haase, Wood and Knops. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Compatibility Between Physical Stimulus Size – Spatial Position and False Recognitions

Seda Dural^{1*}, Birce B. Burhanoğlu¹, Nilsu Ekinci¹, Emre Gürbüz¹, İdil U. Akin¹, Seda Can¹ and Hakan Çetinkaya²

¹ Department of Psychology, İzmir University of Economics, İzmir, Turkey, ² Department of Psychology, Ankara University, Ankara, Turkey

OPEN ACCESS

Edited by:

Krzysztof Cipora,
Universität Tübingen, Germany

Reviewed by:

Jean-Philippe van Dijck,
Ghent University, Belgium
Tina Schmitt,
University of Oldenburg, Germany

*Correspondence:

Seda Dural
seda.dural@ieu.edu.tr

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 17 December 2017

Accepted: 24 July 2018

Published: 14 August 2018

Citation:

Dural S, Burhanoğlu BB, Ekinci N, Gürbüz E, Akin İU, Can S and Çetinkaya H (2018) Compatibility Between Physical Stimulus Size – Spatial Position and False Recognitions. *Front. Psychol.* 9:1457. doi: 10.3389/fpsyg.2018.01457

Magnitude processing is of great interest to researchers because it requires integration of quantity related information in memory regardless of whether the focus is numerical or non-numerical magnitudes. The previous work has suggested an interplay between pre-existing semantic information about number–space relationship in processes of encoding and recall. Investigation of the compatibility between physical stimulus size – spatial position and false recognition may provide valuable information about the cognitive representation of non-numerical magnitudes. Therefore, we applied a false memory procedure to a series of non-numerical stimulus pairs. Three versions of the pairs were used: big-right (a big character on the right/a small character on the left), big-left (a big character on the left/a small character on the right), and equal-sized (an equal sized character on each side). In the first phase, participants ($N = 100$) received 27 pairs, with nine pairs from each experimental condition. In the second phase, nine pairs from each of three stimulus categories were presented: (1) original pairs that were presented in the first phase, (2) mirrored pairs that were horizontally flipped versions of the pairs presented in the first phase, and (3) novel pairs that had not been presented before. The participants were instructed to press “YES” for the pairs that they remembered seeing before and to press “NO” for the pairs that they did not remember from the first phase. The results indicated that the participants made more false-alarm responses by responding “yes” to the pairs with the bigger one on the right. Moreover, they responded to the previously seen figures with the big one on the right faster compared to their distracting counterparts. The study provided evidence for the relationship between stimulus physical size and how they processed spatially by employing a false memory procedure. We offered a size–space compatibility account based on the congruency between the short- and long-term associations which produce local compatibilities. Accordingly, the compatible stimuli in the learning phase might be responsible for the interference, reflecting a possible short-term interference effect on congruency between the short- and long-term associations. Clearly, future research is required to test this speculative position.

Keywords: size–space compatibility, object size, false memory, signal detection, accuracy of recall, reaction time, recall bias

INTRODUCTION

The last two decades have witnessed a flurry of research activity regarding the understanding of extent and nature of number–space association. This activity has been influenced, at least in part, by the work of Dehaene et al. (1990, 1993). They demonstrated that left-hand responses were faster to small as compared to large numbers, whereas the reverse was true for right-hand responses. Moreover, this number–space (response side) compatibility effect was also evident in tasks such as parity judgment that did not require encoding the magnitude of the numbers given. These findings were found to be in line with the metaphor of mental number line (MNL). According to the MNL, numerical representations of magnitudes tend to be spatially organized and the representation of numerical information takes place on an ascending left-to-right oriented line. Based on this, Dehaene et al. (1993) proposed so-called the spatial-numerical association of response codes (SNARC) effect and used the concept of MNL as an account for the SNARC. Since its introduction, several studies have challenged implications of SNARC. Hence, the subsequent research showed that even though the spatial organization of cognitive representations of quantities might be adaptive, the direction and strength of the effect are neither automatic nor unchanging anchors, but are flexible (e.g., Fischer, 2006; Lindemann et al., 2008; Santens and Gevers, 2008; Fischer and Shaki, 2014; Ginsburg et al., 2014; Ginsburg and Gevers, 2015). In addition, according to Proctor and Cho (2006) the endpoints of conceptual dimensions (e.g., tall-short, happy-sad, big-small, etc.) do not share the same representational status, they differ in their valences. People tend to code the stimulus and response alternatives as + polarity and – polarity. Hence, the polarity correspondence account (Proctor and Cho, 2006) predicts that the response selection is faster when the polarities correspond than when they do not. The approach further predicts that the valence of a given pole may be experimentally reversed, because they are largely defined by the relevant context (as in Banks et al., 1975). Although the SNARC effect was largely attributed to representing the numbers along a horizontal line, it may be a consequence of coding large as + polarity and small as – polarity. Therefore, MNL may have originated from ontogenetically acquired behaviors, such as counting (Opfer et al., 2010; Shaki et al., 2012a) or reading and writing habits (Dehaene et al., 1993). As most languages around the world share left-to-right reading and writing direction, MNL appears to be a largely culture-specific, developmentally shaped representational tool which enables efficient coding and comparison of the meaning of magnitudes (Tzelgov et al., 1992).

Magnitude processing is of great interest to researchers because it guides action by integrating information about temporal, spatial, and quantity aspects of the action. Given its significance in survival, the neural mechanism of magnitude processing probably originated from a shared evolutionary history (Hubbard et al., 2005; Cantlon et al., 2009), and thus it might be reasonable to conceptualize a generalized magnitude processing system. In fact, a prominent generalized theory of magnitude (ATOM) (Walsh, 2003, 2015) has already been formulated. According to the theory, information about

time, space, and quantity likely share a common spatial processing mechanism in the brain, due to similarities in their mapping metrics. In line with the theory, the growing body of empirical evidence suggests that a generalized core system may be responsible for the processing of magnitude of different dimensions. The evidence from behavioral studies revealed the relationship between various dimensions, including time, size, letters, luminance (see for reviews Winter et al., 2015; Macnamara et al., 2018), and neurobiological works showed overlapping neural circuits in human parietal cortex for the representation of number, size, and luminance (e.g., Pinel et al., 2001; Fias et al., 2003; Cohen Kadosh et al., 2007; Buetti and Walsh, 2009; Skagerlund et al., 2016).

Although the ATOM hypothesized a general magnitude code serving across diverse quantifiable dimensions, curiously, there has been little work on the relationship between physical size and response location (for a concise review, see Wühr and Seegelke, 2018). Compared to other domains (e.g., number–space and number–size), very few studies (e.g., Ren et al., 2011; Sellaro et al., 2015; Dural et al., 2017; Wühr and Seegelke, 2018) addressed the size and space interaction. For example, in a typical magnitude comparison task, Wühr and Seegelke (2018) found a significant stimulus size–response side compatibility effect when participants were instructed to press left key for the small square, and right key for the large square presented at the center of screen. Participants responded faster to the smaller figure with left key, and faster to the larger figure with right key. They were able to replicate the findings when participants responded to seemingly irrelevant feature (color) of small and large squares. Similarly, Sellaro et al. (2015) asked participants to decide whether the target stimulus was larger or smaller than a reference stimulus with their either right or left index finger. Results revealed a SNARC-like effect: Compared to a reference object, smaller objects were associated with shorter left-side reaction times, and larger objects were associated with shorter right-side reaction times (see Ren et al., 2011; Shaki et al., 2012b, for similar findings). Rather than measuring reaction times, Dural et al. (2017) focused on imagery codes. They presented participants pairs of words referring to objects of varying size differences (e.g., high difference: mouse – elephant, low difference: horse – zebra, average difference: microwave – toaster) and asked to visualize the objects as clearly as possible with eyes closed. After opening their eyes, the participants were instructed to indicate with either left or right hand the location of the imagined objects on the screen divided into halves by a vertical line (e.g., mouse on the left, elephant on the right). Findings showed the tendency to visualize the bigger object on the right increased proportionally with the size difference between the two stimuli, and visualizations of objects seemed to follow an ascending size order from left to right, independent of the hand used to indicate the side of their imagined object. In line with the polarity hypothesis (Proctor and Cho, 2006), the effect tended to diminish as the size difference between the imagined object pairs decreased. These studies provide evidence for a link between mental representations of physical size and space, and this suggestive link manifests itself not only in participants' faster motor responses for the compatible physical stimulus size

and left-right response conditions, but also in how they locate stimuli in space based on relative size. Thus, conceivably, long-term representations play role in physical size and response-side interactions.

In line with the ATOM, successful regulation of action requires integration of quantity related information in memory regardless of whether the focus is numerical or non-numerical magnitudes. The previous work has indicated an interplay between pre-existing semantic information about number-space relationship in processes of encoding and recall. For example, arbitrarily ordered numerical information is not readily stored in the long-term memory; and so, it requires extra effort for acquisition (i.e., training for learning). This working memory account (van Dijck and Fias, 2011) implies that ordinal information is spatially organized not only in long-term memory (Zhang et al., 2016), but also in working memory (Lindemann et al., 2008; Fias et al., 2011; van Dijck and Fias, 2011; Ginsburg and Gevers, 2015). Although recent works have shed light on the role of STM on number-space relationship, to the best of our knowledge, there is only one study (Gut and Staniszewski, 2016) that explicitly addressed the effects of interaction between STM and LTM on the number-space relationship. More specifically, their focus was on how the relatively solid MNL representation modulates the recall of numerical information from STM, regarding the number magnitude-response side congruency. The task they employed required participants to retrieve the spatial position of a digit displayed in the row of four digits which were varied in magnitude. They found that the memorization and retrieval of numbers from STM was more effective when numbers are presented congruently with their position on the LTM.

In cases in which a false memory occurs, participants wrongfully attribute pre-existing semantic information to an external source (Johnson et al., 1993). Thus, memory errors, especially the false alarms, and reaction time measures in recognition may provide helpful data in the understanding of cognitive mechanisms of spatial representations of magnitudes. However, there is so far no evidence of memory influences on relationship between physical size and spatial location of responding and on recall latency and accuracy. If a generalized magnitude coding system is in charge of processing spatially sensitive magnitude information, then it should be possible to identify similar physical size effects on memory performance (e.g., recognition memory) as on numerical magnitude.

In present study, we aimed to investigate the effects of congruency between short- and long-term associations on encoding and retrieval processes in a SNARC-like size-space compatibility by employing a false memory procedure. We manipulated two variables as experimental condition (big-right, big-left, and equal-sized), and stimulus category (original, mirrored, and novel). In the first variable, the big-right and the big-left represented the compatibility and incompatibility conditions, respectively. The equal-sized served as a control condition. In the second variable, as a part of false memory procedure, the original referred to a previously shown stimulus, and the mirrored and novel served as distractors. The study consisted of two main phases as learning and test phases. In the learning phase, we presented a series of non-numerical,

arbitrary pairs of figures, which varied in terms of their relative physical size and spatial position (big-right, big-left, and equal-sized). In the test phase, the participants were tested by original (the same pairs as in learning phase), mirrored (the mirrored versions of the same pairs shown in learning phase), and novel stimulus pairs (the ones never shown in learning phase). The pairs in both learning and test phases always contained identical types of characters. The participants were instructed to indicate as accurately and quickly as possible whether each pair of figures had been seen in a previous phase (i.e., learning phase) of the study. Therefore, the task required comparing the available information (pairs of figures to be tested) with some internal criteria (spatial magnitude representations) that provide guidance on recognition. We evaluated how accurately and how fast participants performed the task.

As may be the case for the numerical comparison tasks, our main prediction is that interaction between memory-dependent information regarding stimulus size and position may interact to elicit a SNARC-like effect. In order to test this, the accuracy and reaction time measures were taken into consideration. We applied Signal Detection Theory to determine discrimination index (d') and response bias (c) based on the observed recognition accuracy in different test conditions. We hypothesize that in comparisons of the previously seen stimulus (original) and the distractors (mirrored and novel), there will be smaller d' values, indicating that participants cannot discriminate signals (previously seen stimuli) from the noise (distractor) when there is compatibility between size and space (i.e., big-right condition). We also expect that the participants will have negative c values in compatible condition, showing a tendency to favor “yes” responses. That is, the semantic map of physical size, which presumably resides in long-term memory, will lead more false recognitions. We also make predictions about the reaction time measures as follows: For the original stimuli, we predict shorter reaction times to the compatible stimuli (big-right) compared to the incompatible stimuli (big-left). As the indicator of interfering effects of the compatible stimuli (big-right) on recall, we predict longer reaction times in distracting conditions (novel and mirrored stimuli). We also predict that the interfering effects in distracting conditions would differ from each other depending on whether the distractors consist of novel stimuli or altered versions of the originally seen stimuli.

MATERIALS AND METHODS

Participants

A total of 100 participants (39 males and 61 females) took part in the experiment. They were university students and staff, aged between 19 and 32 years ($M = 22.56$, $SD = 2.04$). All participants were right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971; $LQ > +50$), had normal or corrected-to-normal vision, and no history of neurological or psychiatric disorders. They gave written informed consent in accordance with the ethics committee of the Izmir University of Economics (B.30.2.IEU.0.05.05-020-054), where the study was carried out.

Stimuli

Thirty-six characters (Ω, Ξ, Δ, Σ, Ι, Ο, Λ, Ε, Ζ, Η, Π, Σ, Ω, Γ, Λ, Μ, Ψ, Ζ, Η, Κ, Υ, C, S, Γ, Ο, Δ, S, Γ, Ψ, X, S, Θ, Ξ, Δ, Ε) were obtained from Microsoft Word symbols. Ten were symmetrical (e.g., Ω, Λ, S) and 26 were asymmetrical (e.g., Ι, Ο, Σ, Η). These characters were assigned to one of three experimental conditions (big-right, big-left, and equal-sized). All but equal-sized condition consisted of symmetrical and asymmetrical characters. In equal-sized condition, we used only asymmetrical characters in order to create proper testing stimuli for the mirrored condition (since the mirrored images of the symmetrical characters would not be proper distractors). Therefore, 12 out of the 26 asymmetrical characters were randomly selected for the equal-sized condition. Then, an equal number of characters were randomly assigned to the big-right and the big-left conditions, chosen from the remaining 24.

These characters were used to construct stimulus pairs. Each pair consisted of two identical characters that varied in size depending on the experimental condition. These were presented on a 19.5-inch LCD display at full 1600 × 900 pixel resolution 40 cm away from the participant, which corresponds to 48° × 32° of visual angle. The pairs of characters were vertically centered and positioned 400 pixels away from each side of the screen. Each character was presented in an imaginary square placeholder. The angular sizes of figures were 8.53° × 8.53° for the larger versions, 2.15° × 2.15° for the smaller versions, and 4.29° × 4.29° for the equal-sized versions. All characters were in black, with a white background.

Three different types of stimulus pairs comprised the experimental conditions. A big-right pair was constructed with a big version of the character on the right, and a small version of the same character on the left. A big-left pair contained a big version on the left and a small version of the same character on the right. For an equal-sized pair, an equal-sized version appeared on each side. Thus, there was a total of 36 pairs (12 big-right pairs, 12 big-left pairs, and 12 equal-sized pairs).

Nine out of the 12 big-right pairs construed the big-right condition of the learning phase (**Figure 1A**). The remaining three pairs functioned as novel stimuli in the test phase (**Figure 1B**). Three of the nine big-right pairs used in the learning phase functioned as original stimuli in the test phase, and another three different pairs (i.e., not including original stimulus) out of the nine learning pairs functioned as mirrored stimuli in the test phase (**Figures 1A,B**). Mirrored pairs were constructed by horizontally flipping the individual characters and their spatial positions (i.e., left or right). The constructed pairs were randomly assigned to the conditions, and the same procedure was followed for big-left and equal-sized pairs.

Procedure

The participants were seated comfortably in a dimly lit sound-attenuating chamber, and were instructed to remain in the same position throughout the experiment. The experiment was carried out in two successive phases, with a filler task between (**Figure 2**). In the learning phase (**Figure 2A**), the participants were presented a total of 27 stimulus pairs, which consisted of nine pairs from each experimental condition (big-right, big-left,

and equal-sized). They were instructed to memorize as many pairs as possible, by considering the form, size, and spatial location of the stimuli on the screen. At the end of the learning phase, a brief filler task was introduced to prevent any rehearsal (**Figure 2B**). The filler task consisted of 10 simple arithmetic calculations¹ [e.g., $(76 \div 2) \times 4$ and $(979 - 779) \div 2$].

A total of 27 stimulus pairs were presented in the test phase (**Figure 2C**): nine in the original stimulus category consisting of an equal number of big-right, big-left, and equal-sized pairs; nine in the mirrored stimulus category consisting of an equal number of big-right, big-left, and equal-sized pairs; and nine in novel stimulus category consisting of an equal number of big-right, big-left, and equal-sized pairs. Each pair was presented in a randomized order for 2,000 ms, with a 500 ms inter-stimulus interval both in the learning and test phases. The participants were asked to indicate whether they had previously seen that specific pair of characters in the learning phase by pressing B key for a “YES” response or N key for a “NO” response as quickly as possible. The participants used their right index finger for the “YES” response and their right middle finger for the “NO” response. Their responses yielded two measures, accuracy of recall and reaction times. SuperLab 4.5² (Cedrus Corporation, United States) was used to control stimulus presentations and response recordings during the experimental sessions. It took about 10 min for each participant to complete the task.

Data Analysis

Experimental condition (big-right, big-left, and equal-sized) and stimulus category (original, mirrored, and novel) were within-participant variables. Accuracy of recall and reaction time were recorded as dependent measures. Response accuracy was examined within the framework of signal detection theory. In addition, for each condition, mean reaction time scores were calculated disregarding the accuracy of responses³. A 3 (experimental condition) × 3 (stimulus category) repeated analysis of variance (ANOVA) was conducted to analyze the reaction time data. In the analysis of reaction time data, in all planned contrasts, the original stimulus category was used as the reference condition for the stimulus category, and the big-right condition was used as the reference condition for the experimental condition.

Signal Detection Analysis

Signal detection theory is an accepted procedure when signal and noise trials must be discriminated (Stanislaw and Todorov, 1999). In this study, we define signal trials as those that contain

¹Arithmetic problems are commonly presented as filler task in false memory studies (e.g., Coane and McBride, 2006). In order to ensure completeness, we checked the answers for their accuracy, and found that the arithmetic problems were 85–90% percent solved with accuracy.

²<http://www.superlab.com/>

³In analyses of repeated measures, when we obtain a “0” accuracy score from a participant in a specific condition (e.g., mirrored stimulus category), it is not possible to use other responses of the same participant to calculate the model parameters. Further, when the reaction time data for only correct responses were analyzed, the results maintained for the original main and interaction effects (see the **Supplementary Material**). Thus, the reaction time data for both correct and incorrect responses were reported.





































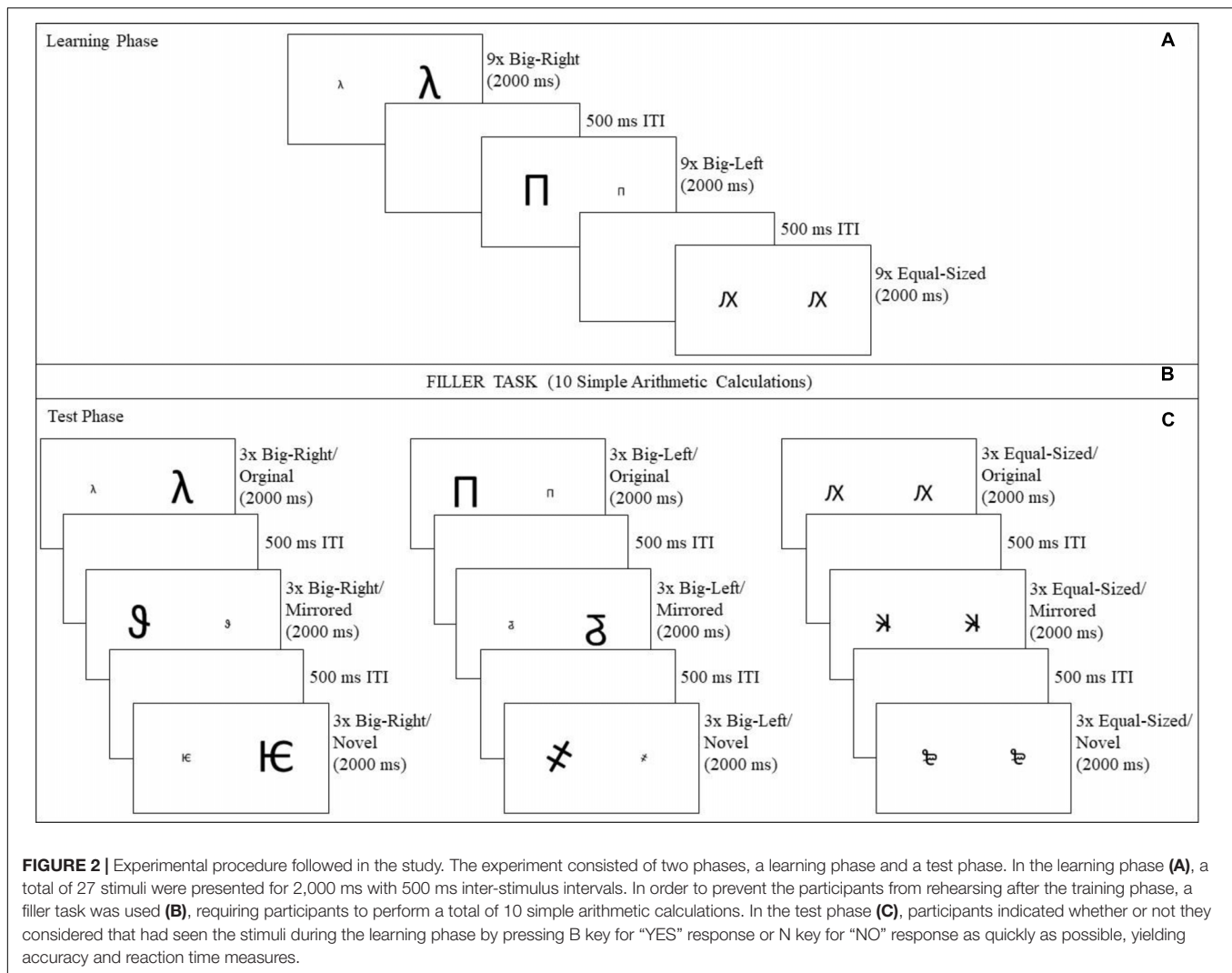
A Learning Phase		B Test Phase	
Big-Right (x9)	 	Original (x3)	 
	 	Mirrored (x3)	 
	 	Novel (x3)	 
Big-Left (x9)	 	Original (x3)	 
	 	Mirrored (x3)	 
	 	Novel (x3)	 
Equal-Sized (x9)	 	Original (x3)	 
	 	Mirrored (x3)	 
	 	Novel (x3)	 

FIGURE 1 | Stimulus pairs used in the learning and test phases of the experiment. In the learning phase **(A)** a total of 27 stimulus pairs with nine stimulus pairs from each experimental condition (big-right, big-left, and equal-sized) were presented in random order. In the test phase **(B)**, nine stimulus pairs (3 × each experimental condition) from the learning phase were used as the original stimuli, nine mirrored version of the stimulus pairs (3 × each experimental condition) from the learning phase, as the mirrored stimuli, and nine new stimulus pairs (3 × each experimental condition), as the novel stimuli.



previously studied stimuli, and noise trials as those that contain distractor stimuli of yes/no task (e.g., seen/unseen). On signal trials, “yes” responses are correct and are named as *hit*. On noise trials, however, “yes” responses are incorrect and are termed as *false alarm*. The hit rate (the probability of responding yes on signal trials) and the false alarm rate (the probability of responding yes on noise trials) are the indicators of performance in a yes/no task. The hit rate is calculated by dividing the number of hits by the total number of signal trials. The false alarm rate is calculated by dividing the number of false alarms by the total number of noise trials. Based on these hit and false-alarm rates, two signal detection parameters are calculated: sensitivity (d') and response bias (c).

d' represents the participants' ability to discriminate the “signals” (hits) from the “noise” (false alarms) (Wilson and Swets, 1954). This is calculated by subtracting the z -score of false-alarm rate from the z -score of hit rate. A d' value of 0 (zero) indicates an inability to distinguish signal from noise, whereas higher values reflect more “yes” responses to previously studied stimuli, and more “no” responses to distracting stimuli (Lockhart

and Murdock, 1970). c is calculated by averaging the z scores of hit and false alarm rates, then multiplying the result by -1 . Negative values of c indicate a bias toward “yes” responses, and positive values, in favor of “no” responses (Stanislaw and Todorov, 1999).

Accordingly, in the present study, stimulus pairs of the original stimulus category were identified as the signal, and stimulus pairs of novel and mirrored stimulus categories, as the noise. Thus, “yes” responses in the original stimulus category constituted hits, and “yes” responses in the novel and mirrored stimulus categories, false alarms. In regard to the experimental conditions, hit and false alarm values were calculated in six parts (Table 1). d' and c parameters for each participant were calculated based on these parts. For example, to calculate d' and c values in the original versus novel comparison of the big-right condition (see row 1/ Table 1), signal trials were acquired from the big-right/original stimuli, and hits were gathered from “yes” responses to those stimuli. For the noise trials, big-right/novel stimuli were used, and false alarms were obtained from the “yes” responses to those stimuli.

TABLE 1 | Stimulus category comparisons by experimental conditions used for calculating hit and false alarm rates.

		Previously studied (hits)	Distractor (false alarms)
Original versus novel	1	Big-right/original	Big-right/novel
	2	Big-left/original	Big-left/novel
	3	Equal-sized/original	Equal-sized/novel
Original versus mirrored	4	Big-right/original	Big-right/mirrored
	5	Big-left/original	Big-left/mirrored
	6	Equal-sized/original	Equal-sized/mirrored

TABLE 2 | Mean and standard deviation values of d' and c parameters in the original versus novel and the original versus mirrored comparisons by experimental conditions.

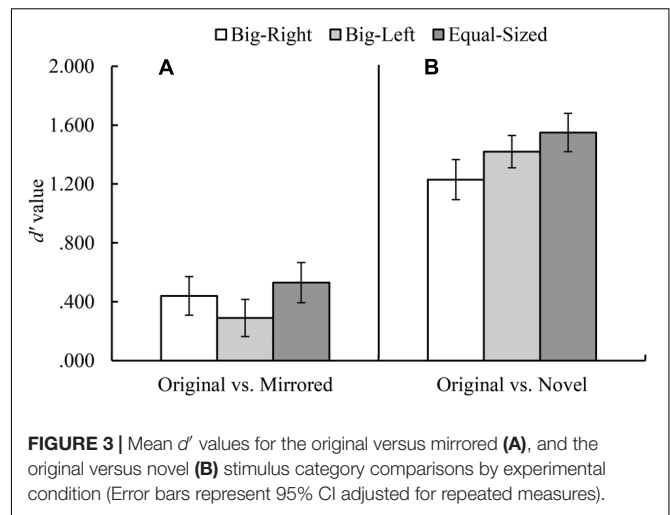
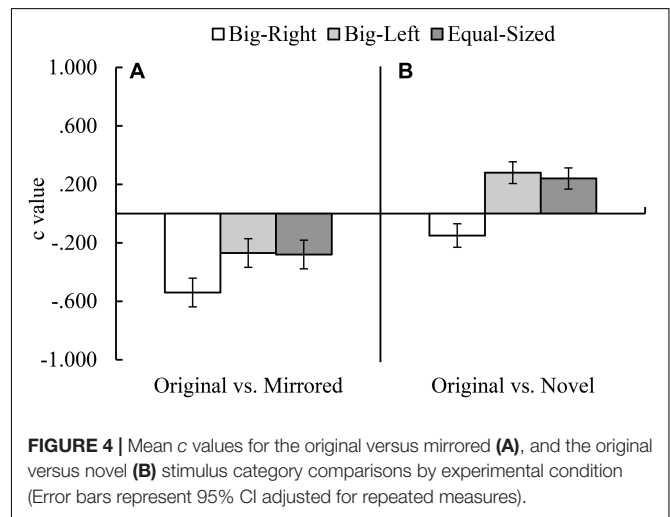
	Experimental condition					
	Big-right		Big-left		Equal-sized	
	d'	c	d'	c	d'	c
Original versus novel	1.229 (0.861)	-0.145 (0.407)	1.415 (0.638)	0.279 (0.381)	1.546 (0.687)	0.242 (0.363)
Original versus mirrored	0.440 (0.771)	-0.538 (0.498)	0.293 (0.859)	-0.272 (0.497)	0.528 (0.794)	-0.283 (0.500)

RESULTS

Accuracy of Recall

Four separate one-way repeated ANOVAs were performed both in the original versus mirrored comparison and the original versus novel comparison for d' and c parameters. In the analysis of the signal detection parameters, the big-right condition was used as the reference condition in all planned contrasts. **Table 2** shows mean and standard deviation values of d' and c parameters in the original versus novel, and the original versus mirrored comparisons by experimental conditions.

In the original versus novel stimulus category comparison for d' parameter, Mauchly's test indicated that assumption of sphericity had been violated, $\chi^2_{(2)} = 10.59$, $p = 0.005$. Therefore, degrees of freedom were corrected by using Greenhouse–Geisser estimates of sphericity. The results of the analysis indicated a significant experimental condition effect for d' parameter, $F_{(1.81, 177.62)} = 5.23$, $p = 0.008$, $\eta^2 = 0.05$ (**Figure 3B**). Contrasts based on d' values indicated that the participants in the big-right condition performed worse than those in the equal-sized condition, $F_{(1,98)} = 8.08$, $p = 0.005$, $\eta^2 = 0.08$, in discriminating the signal from noise; but they performed similarly in the big-left condition, $F_{(1,98)} = 3.59$, $p = 0.061$. In the original versus novel stimulus category comparison for c parameter, it was found a significant effect of experimental condition, $F_{(2,194)} = 49.46$, $p < 0.001$, $\eta^2 = 0.34$ (**Figure 4B**). Contrasts analysis based on c values revealed that the participants significantly favored the “yes” response in the big-right condition compared to the big-left condition, $F_{(1,97)} = 81.33$, $p < 0.001$, $\eta^2 = 0.47$, and the equal-sized condition, $F_{(1,97)} = 67.04$, $p < 0.001$, $\eta^2 = 0.41$.

**FIGURE 3 |** Mean d' values for the original versus mirrored (A), and the original versus novel (B) stimulus category comparisons by experimental condition (Error bars represent 95% CI adjusted for repeated measures).**FIGURE 4 |** Mean c values for the original versus mirrored (A), and the original versus novel (B) stimulus category comparisons by experimental condition (Error bars represent 95% CI adjusted for repeated measures).

In the original versus mirrored stimulus category comparison, d' values did not differ across the experimental conditions, $F_{(2,198)} = 2.53$, $p = 0.082$ (**Figure 3A**); on the other hand, c values indicated a significant effect of the experimental condition, $F_{(2,198)} = 11.65$, $p < 0.001$, $\eta^2 = 0.11$ (**Figure 4A**). Planned contrasts based on the c values indicated that the participants favored the “yes” response more in the big-right condition compared to both the big-left, $F_{(1,99)} = 20.42$, $p < 0.001$, $\eta^2 = 0.17$, and equal-sized conditions, $F_{(1,99)} = 15.36$, $p < 0.001$, $\eta^2 = 0.13$.

Reaction Time

The mean and standard deviation values of reaction time scores for the experimental conditions by stimulus categories are shown in **Table 3**. Mauchly's test indicated that assumption of sphericity had been violated for the main effect of experimental condition, $\chi^2_{(2)} = 12.37$, $p = 0.002$ and for the interaction between experimental condition and stimulus category, $\chi^2_{(9)} = 22.71$, $p = 0.007$. Therefore, degrees of freedom were corrected by using Greenhouse–Geisser estimates of sphericity. A 3×3 ANOVA

TABLE 3 | Mean and standard deviation values of reaction time scores in experimental conditions by stimulus categories.

		Experimental condition		
		Big-right	Big-left	Equal-sized
Stimulus category	Original	1262.105 (387.928)	1431.962 (465.016)	1259.870 (323.795)
	Novel	1538.702 (594.657)	1158.742 (316.959)	1168.950 (327.136)
	Mirrored	1680.182 (701.887)	1477.397 (508.622)	1383.448 (427.413)

for repeated measures indicated a significant main effect of the experimental condition on reaction time, $F_{(1.79, 177.02)} = 22.59$, $p < 0.001$, $\eta^2 = 0.19$. Planned contrasts revealed that the reaction time for the big-right ($mean = 1483.66$, $SE = 44.36$) condition was significantly longer than the big-left ($mean = 1356.03$, $SE = 32.35$), $F_{(1,99)} = 15.79$, $p < 0.001$, $\eta^2 = 0.14$, and equal-sized ($mean = 1270.76$, $SE = 25.80$) conditions, $F_{(1,99)} = 35.32$, $p < 0.001$, $\eta^2 = 0.26$. There was also a significant main effect of the stimulus category on reaction time, $F_{(2,198)} = 25.71$, $p < 0.001$, $\eta^2 = 0.21$. The contrasts analysis indicated that the mean reaction time for the original stimulus category ($mean = 1317.98$, $SE = 30.36$) was significantly shorter than that for the mirrored stimulus category ($mean = 1513.68$, $SE = 42.21$), $F_{(1,99)} = 36.17$, $p < 0.001$, $\eta^2 = 0.27$; however, it did not differ from the mean reaction time for the novel stimulus category ($mean = 1288.80$, $SE = 31.93$), $F_{(1,99)} = 0.78$, $p = 0.380$.

There was a significant interaction effect between the experimental condition and the stimulus category, $F_{(3.61, 356.96)} = 16.21$, $p < 0.001$, $\eta^2 = 0.14$. Four planned contrasts were performed comparing each level of stimulus categories (i.e., novel and mirrored) to the original stimulus category across each level of experimental conditions (i.e., big-left and equal-sized) comparing to the big-right condition. The first contrast that compared the original stimulus category to the novel stimulus category in respect to the big-right and big-left conditions was significant, $F_{(1,99)} = 50.58$, $p < 0.001$, $\eta^2 = 0.34$. This significant interaction indicated that the participants responded faster to the big-right than to the big-left stimuli in the original stimulus category; however, in the novel stimulus category, reaction times were slower on the big-right compared to the big-left condition (**Figure 5A**). The second contrast was performed to compare the reaction time data obtained from the original stimulus category and from the novel stimulus category, in respect to the big-right and equal-sized conditions. This interaction was also significant, $F_{(1,99)} = 30.44$, $p < 0.001$, $\eta^2 = 0.24$, suggesting that reaction times were similar for both the big-right and equal-sized conditions in the original stimulus category; they were slower for the big-right condition than the equal-sized condition in the novel stimulus category (**Figure 5B**). The third contrast which compared the original and mirrored stimulus categories in respect to the big-right and big-left conditions was significant, $F_{(1,99)} = 24.79$, $p < 0.001$, $\eta^2 = 0.20$. This significant interaction suggested that participants responded

faster to the big-right stimuli than to the big-left stimuli in the original stimulus category; but in the mirrored stimulus category, they responded more slowly to the big-right stimuli than to the big-left stimuli (**Figure 5C**). The final contrast, which compared the original and mirrored stimulus categories in respect to the big-right and equal-sized conditions, was also significant, $F_{(1,99)} = 14.31$, $p < 0.001$, $\eta^2 = 0.13$. This significant interaction implied that the reaction times obtained from the big-right and equal-sized conditions on the original stimulus category were similar; however, they were longer on the big-right condition than the equal-sized condition in the mirrored stimulus category (**Figure 5D**).

DISCUSSION

The aim of the present study was to see whether we could elicit any evidence for an association between size and space by investigating the recognition memory. Particularly, we predicted that compatibility between physical size – spatial position of stimulus and its memory representation would affect the sensitivity, response bias, and reaction time for recognition. Hence, the big-right (also, small-left) stimuli appeared to be associated significantly faster responses when tested with original pairs of stimuli. Whereas, big-right distractors (i.e., novel and mirrored trials) produced increased number of false “yes” responses and longer reaction time measures. This implies a compatibility effect between size and space regarding memory representations of size and horizontal position of stimuli. Thus, an important implication of this finding pertains to ATOM (“A Theory Of Magnitude”; Walsh, 2003, 2015; Buetti and Walsh, 2009) which predicts the existence of a generalized, integrating magnitude-processing system that helps the control of complex actions by providing a ground for interacting of magnitudes such as number, space, and time.

We suggest from our results that the system might be activated upon the detection of a size difference between the two characters. Compared to the other experimental conditions, the equal-sized condition yielded better discrimination, lower response bias, and similar reaction times regardless of the stimulus category tested with (whether or not the test stimulus was original, mirrored, or novel). The equal-sized condition probably would not activate the magnitude comparing process, because it would have been redundant, presumably the ATOM operates on the size–space compatibility. Similarly, when there was size difference in a pair of characters (i.e., big-left and big-right conditions), we observed differences in reaction time and the signal detection parameters. Therefore, we conceived that this generalized magnitude processing system might be sensitive to the size differences for its activation: An internal on-off switch might be operated by a perceptual process upon the detection of a size difference.

However, ATOM does not assume the direction of the size–spatial position association readily (Walsh, 2003, 2015; Buetti and Walsh, 2009). Then, this left-to-right oriented size–space compatibility effect we observed, calls for an explanation. Several accounts may be offered. Some recent data has indicated

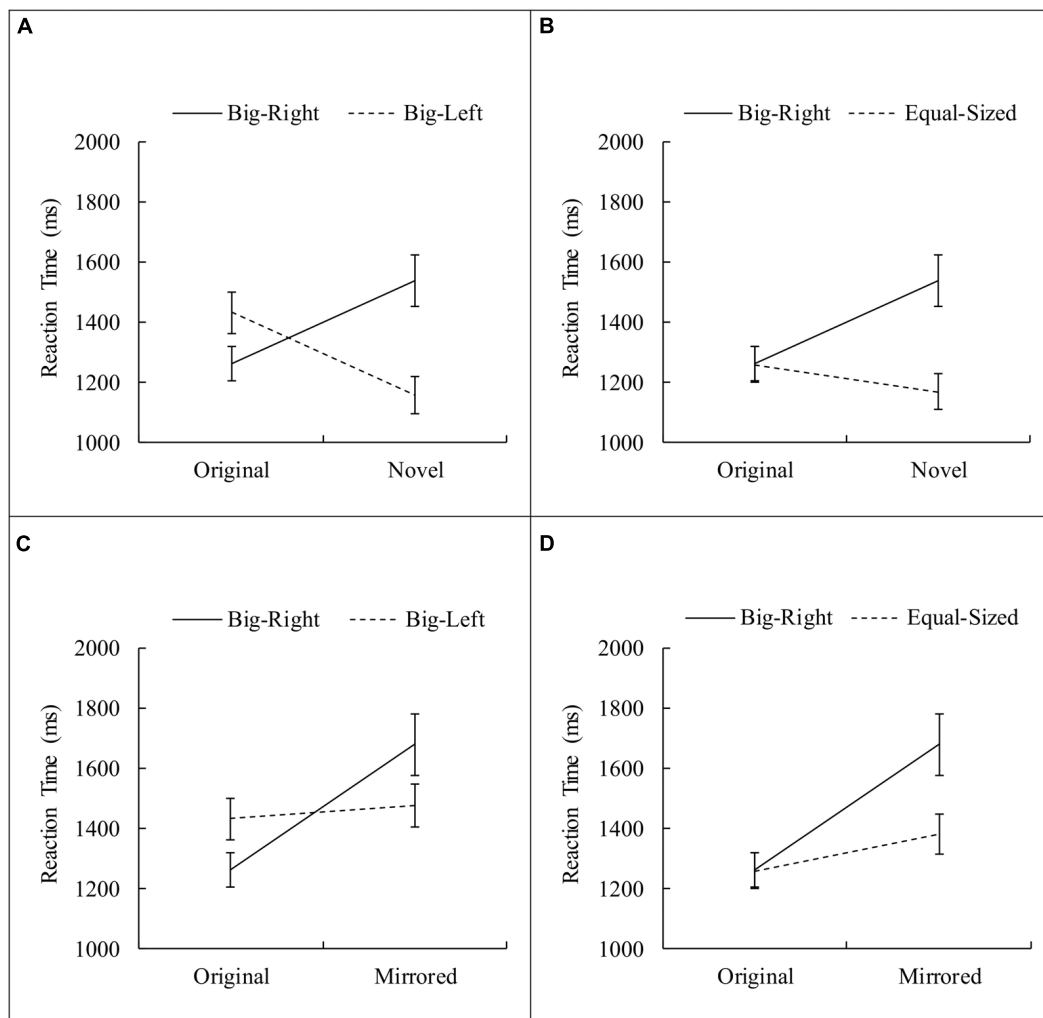


FIGURE 5 | Planned contrasts for the interaction between experimental condition and stimulus category. **(A)** Shows data for original vs. novel and big-right vs. big-left; **(B)** shows data for original vs. novel and big-right vs. equal-sized; **(C)** shows data for original vs. mirrored and big-right vs. big-left; **(D)** shows data for original vs. mirrored and big-right vs. equal-sized (Error bars represent 95% CI adjusted for repeated measures).

functional differences between the two hands. The right hand generally is the dominant hand and it is stronger than the left hand (Hepping et al., 2015). Hence, perhaps the manual-motor dominance gives way to the faster right-hand responses to the larger magnitudes (Incel et al., 2002). In our study, our participants were all right-handers, and they used only their right hand to respond to the stimuli. Moreover, they responded to the stimuli by hitting the “b” or “n” keys which are located centrally on a QWERTY keyboard. Therefore, we do not consider this as a valid account for our experimental setup. On the other hand, if the compatibility between physical size and space utilizes the same sources as the SNARC, then we may explain the direction of the effect based on the converging support of acculturation such as diffusion of spatial-directional scanning habits from reading into the domain of numerical (Dehaene et al., 1993) or other magnitude-related cognition, consensually developed action patterns (Lindemann et al., 2011), and the influence of

external representations such as graphs and notation systems (Bender et al., 2010; Tversky, 2011).

Alternatively, given the methodological differences between our study and the previous work, we offered an account based on congruency between the short-term and long-term associations. This account adapted from dual-route models (Tagliabue et al., 2000) by Wühr and Seegeler (2018) to explain how relevant and irrelevant stimulus features evoke short-term and long-term associations. The account predicts that when there is congruency between these two associations, both processing routes activate the correct response resulting in shorter reaction times and better accuracy. Whereas in incongruent conditions, the long-term association would interfere with selection of the correct response resulting in longer reaction times and lessened accuracy. Similarly, we assume that the information about the spatial orientation of magnitudes has already been stored in memory of long-term association between size and space, probably through

TABLE 4 | Congruency between short- and long-term associations in big-right and big-left conditions by different stimulus categories.

	Experimental condition	Stimulus category	Short-term	Long-term	Congruency
1	Big-right	Original	Compatible	Compatible	Congruent
2	Big-left	Original	Compatible	Incompatible	Incongruent
3	Big-right	Novel	Incompatible	Compatible	Incongruent
4	Big-left	Novel	Incompatible	Incompatible	Congruent
5	Big-right	Mirrored	Incompatible	Incompatible	Congruent
6	Big-left	Mirrored	Incompatible	Compatible	Incongruent

acculturation. In fact, according to the instances theory (Choplin and Logan, 2005), our past experiences are the main source for the magnitude–space associations, thus people rely on the instances available from long-term memory (Logan, 1988). Likewise, this congruency account could be applied to size–space compatibility effect that we observed in our study.

We conceived that short-term associations reflect the decisions of participants about the test stimuli based on acquired size – spatial position information through the learning phase of the study. The level of the short-term association therefore depends, in part, on the compatibility between the test stimuli and the stimuli studied in learning phase. Whereas, the long-term associations refer to the information about the size – spatial position had already been acquired. The level of the long-term association depends on the compatibility between the test stimuli and the available information acquired through long-term processing. In the original stimulus category, the test stimuli were the same as the stimuli presented during the learning phase. Here, participants showed shorter reaction times to the big-right stimuli, compared to the big-left stimuli. Given the fact that both the big-right and big-left stimuli were tested by their exact copies (short-term compatibility; see **Table 4**, rows 1 and 2), the difference observed in reaction times may be attributed to the differential long-term compatibility levels of the big-right (long-term compatibility; see **Table 4**, row 1) and big-left (long-term incompatibility; see **Table 4**, row 2) stimuli. Hence, the congruency between short- and long-term associations in the big-right condition resulted in decreased reaction times.

This congruency effect was also evident in the novel condition in which participants were tested with stimuli that they did not studied in learning phase. Interestingly, this time the culprit was the big-left stimuli: Being as the novel stimuli, both the big-right and big-left stimuli were not encoded in the learning phase (short-term incompatibility; **Table 4**, rows 3 and 4); however, the big-left stimuli were structurally incompatible with the long-term association code (long-term incompatibility; **Table 4**, row 4). This reflects a negative congruency between the short- and long-term associations in the big-left condition. Hence, we obtained shorter reaction times and lower false alarm rates. On the other hand, in the big-right stimulus category, the test stimuli were compatible with the long-term association code (long-term compatibility; **Table 4**, row 3). This reflects an incongruency between the short- and long-term associations in the big-right condition. As a result, we observed an interference on decisions of participants presumably originated from the long-term association code.

This interference was evidenced by longer reaction times and higher false alarm rates. This indicates a long-term interference effect on congruency between the short- and long-term associations.

Finally, in the mirrored stimulus category, participants were tested with mirrored (horizontally flipped) versions of the stimuli. Being as the mirrored stimuli, both the big-right and big-left stimuli were not the same as what had been seen in learning phase (short-term incompatibility; **Table 4**, rows 5 and 6). However, the big-right mirrored stimuli were structurally incompatible with the long-term association code (long-term incompatibility; **Table 4**, row 5). Hence, we observed another negative congruency effect in the mirrored condition with the big-right stimuli. However, this time, the observed congruency resulted in increased reaction times and increased response bias in favor of “yes.” This is an unexpected finding, because instead of the expected facilitating effect of the congruency, we obtained an interference. Obviously, in the big-right mirrored trials, the long-term compatible big-right stimuli seen in the learning phase were tested with the long-term incompatible stimuli. We speculate, therefore, that the compatible stimuli in the learning phase might be responsible for the interference, reflecting a possible short-term interference effect on congruency between the short- and long-term associations. Clearly, future research is required to test this speculative position.

CONCLUSION

To conclude, we provided evidence for the relationship between stimulus physical size and how they processed spatially by employing a false memory procedure. To the best of our knowledge, this is the first work that uses memory errors to investigate the size–space relationship. Also, this piece of evidence supported the existence of a generalized magnitude processing system assumed by ATOM. Since the ATOM lacks an account of the direction of the size – spatial position association, we offered an interplay between the short-term and long-term associations which determines the direction of the spatial organization of physical magnitudes. Thus, in line with Ginsburg and Gevers (2015) and Gut and Staniszcwski (2016), spatial response biases might result from the activation of both pre-existing positions and from temporary space associations at the same time. Finally, we offer a size–space compatibility account based on the congruency between the short- and long-term associations which produce local compatibilities.

We think that our study takes place in the intersection of shared-representation and shared-decision accounts and offers more eclectic approach toward the understanding of magnitude–space association. Future research is required to further test the suggestive evidence provided by the present study.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

- Banks, W. P., Clark, H. H., and Lucy, P. (1975). The locus of the semantic congruity effect in comparative judgments. *J. Exp. Psychol. Hum. Percept. Perform.* 104, 35–47. doi: 10.1037/0096-1523.1.1.35
- Bender, A., Hutchins, E., and Medin, D. (2010). Anthropology in cognitive science. *Top. Cogn. Sci.* 2, 374–385. doi: 10.1111/j.1756-8765.2010.01082.x
- Bueti, D., and Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1831–1840. doi: 10.1098/rstb.2009.0028
- Cantlon, J. F., Platt, M. L., and Brannon, E. M. (2009). Beyond the number domain. *Trends Cogn. Sci.* 13, 83–91. doi: 10.1016/j.tics.2008.11.007
- Choplin, J. M., and Logan, G. D. (2005). A memory-based account of automatic numerosity processing. *Mem. Cognit.* 3, 17–28. doi: 10.3758/BF03195294
- Coane, J. H., and McBride, D. M. (2006). The role of test structure in creating false memories. *Mem. Cognit.* 34, 1026–1036. doi: 10.3758/BF03193249
- Cohen Kadosh, R., Cohen Kadosh, K., Linden, D. J., Gevers, W., Berger, A., and Henik, A. (2007). The brain locus of interaction between number and size: a combined functional magnetic resonance imaging and event-related potential study. *J. Cogn. Neurosci.* 19, 957–970. doi: 10.1162/jocn.2007.19.6.957
- Dehaene, S., Bossini, S., and Giraux, P. (1993). The mental representation of parity and number magnitude. *J. Exp. Psychol. Gen.* 122, 371–396. doi: 10.1037/0096-3445.122.3.371
- Dehaene, S., Dupoux, E., and Mehler, J. (1990). Is numerical comparison digital? Analogical and symbolic effects in two-digit number comparison. *J. Exp. Psychol. Hum. Percept. Perform.* 16, 626–641. doi: 10.1037/0096-1523.16.3.626
- Dural, S., Çetinkaya, H., and Güntürkün, O. (2017). Imagine a mouse and an elephant: hemispheric asymmetries of imagination. *Laterality Asymmetries Body Brain Cogn.* 22, 354–361. doi: 10.1080/1357650X.2016.1200594
- Fias, W., Lammertyn, J., Reynvoet, B., Dupont, P., and Orban, G. A. (2003). Parietal representation of symbolic and non-symbolic magnitude. *J. Cogn. Neurosci.* 15, 47–56. doi: 10.1162/089892903321107819
- Fias, W., van Dijck, J. P., and Gevers, W. (2011). “How is number associated with space? The role of working memory,” in *Space, Time and Number in the Brain Searching for the Foundations of Mathematical Thought*, eds S. Dehaene and E. Brannon (New York, NY: Elsevier), 133–148. doi: 10.1016/B978-0-12-385948-8.00010-4
- Fischer, M. H. (2006). The future for SNARC could be star. *Cortex* 42, 1066–1068. doi: 10.1016/S0010-9452(08)70218-1
- Fischer, M. H., and Shaki, S. (2014). Spatial associations in numerical cognition—From single digits to arithmetic. *Q. J. Exp. Psychol.* 67, 1461–1483. doi: 10.1080/17470218.2014.927515
- Ginsburg, V., Dijck, J. V., Previtali, P., Fias, W., and Gevers, W. (2014). The impact of verbal working memory on number–space associations. *J. Exp. Psychol. Learn. Mem. Cogn.* 40, 976–986. doi: 10.1037/a0036378
- Ginsburg, V., and Gevers, W. (2015). Spatial coding of ordinal information in short- and long-term memory. *Front. Hum. Neurosci.* 9:8. doi: 10.3389/fnhum.2015.00008
- Gut, M., and Staniszevski, R. (2016). The SNARC effect in number memorization and retrieval. What is the impact of congruency, magnitude and the exact position of numbers in short-term memory processing? *Adv. Cogn. Psychol.* 12, 193–208. doi: 10.5709/acp-0198-0
- Hepping, A. M., Ploegmakers, J. J., Geertzen, J. H., Bulstra, S. K., and Stevens, M. (2015). The Influence of hand preference on grip strength in children and adolescents; A cross-sectional study of 2284 children and adolescents. *PLoS One* 10:e0143476. doi: 10.1371/journal.pone.0143476
- Hubbard, E. M., Piazza, M., Pinel, P., and Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nat. Rev. Neurosci.* 6, 435–448. doi: 10.1038/nrn1684
- Incel, N. A., Ceceli, E., Durukan, P. B., Erdem, H. R., and Yorgancioglu, Z. R. (2002). Grip strength: effect of hand dominance. *Singapore Med. J.* 43, 234–237.
- Johnson, M. K., Hashtroudi, S., and Lindsay, D. S. (1993). Source monitoring. *Psychol. Bull.* 114, 3–28. doi: 10.1037/0033-2909.114.1.3
- Lindemann, O., Abolafia, J. M., Pratt, J., and Bekkering, H. (2008). Coding strategies in number space: memory requirements influence spatial-numerical associations. *Q. J. Exp. Psychol.* 61, 515–524. doi: 10.1080/17470210701728677
- Lindemann, O., Alipour, A., and Fischer, M. H. (2011). Finger counting habits in middle-eastern and western individuals: an online survey. *J. Cross Cult. Psychol.* 42, 566–578. doi: 10.1177/0022022111406254
- Lockhart, R. S., and Murdock, B. B. Jr. (1970). Memory and the theory of signal detection. *Psychol. Bull.* 74, 100–109. doi: 10.1037/h0029536
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychol. Rev.* 95, 492–527. doi: 10.1037/0033-295X.95.4.492
- Macnamara, A., Keage, H. A. D., and Loetscher, T. (2018). Mapping of non-numerical domains on space: a systematic review and meta-analysis. *Exp. Brain Res.* 236, 335–346. doi: 10.1007/s00221-017-5154-6
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4
- Opfer, J. E., Thompson, C. A., and Furlong, E. E. (2010). Early development of spatial-numeric associations: evidence from spatial and quantitative performance of preschoolers. *Dev. Sci.* 13, 761–771. doi: 10.1111/j.1467-7687.2009.00934.x
- Pinel, P., Dehaene, S., Rivière, D., and LeBihan, D. (2001). Modulation of parietal activation by semantic distance in a number comparison task. *Neuroimage* 14, 1013–1026. doi: 10.1006/nimg.2001.0913
- Proctor, R. W., and Cho, Y. S. (2006). Polarity correspondence: a general principle for performance of speeded binary classification tasks. *Psychol. Bull.* 132, 416–442. doi: 10.1037/0033-2909.132.3.416
- Ren, P., Nicholls, M. R., Ma, Y., and Chen, L. (2011). Size matters: non-numerical magnitude affects the spatial coding of response. *PLoS One* 6:e23553. doi: 10.1371/journal.pone.0023553
- Santens, S., and Gevers, W. (2008). The SNARC effect does not imply a mental number line. *Cognition* 108, 263–270. doi: 10.1016/j.cognition.2008.01.002
- Sellaro, R., Treccani, B., Job, R., and Cubelli, R. (2015). Spatial coding of object typical size: evidence for a SNARC-like effect. *Psychol. Res.* 79, 950–962. doi: 10.1007/s00426-014-0636-7
- Shaki, S., Fischer, M. H., and Göbel, S. M. (2012a). Direction counts: a comparative study of spatially directional counting biases in cultures with different reading directions. *J. Exp. Child Psychol.* 112, 275–281. doi: 10.1016/j.jecp.2011.12.005

ACKNOWLEDGMENTS

We thank Prof. Michael Domjan of University of Texas at Austin for proofreading the manuscript and Prof. Harold Stanislaw of California State University, Stanislaus for helpful discussion and advice regarding the signal detection analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.01457/full#supplementary-material>

- Shaki, S., Petrusic, W. M., and Leth-Steensen, C. (2012b). SNARC effects with numerical and non-numerical symbolic comparative judgments: instructional and cultural dependencies. *J. Exp. Psychol. Hum. Percept. Perform.* 38, 515–530. doi: 10.1037/a0026729
- Skagerlund, K., Karlsson, T., and Träff, U. (2016). Magnitude processing in the brain: an fMRI study of time, space, and numerosity as a shared cortical system. *Front. Hum. Neurosci.* 10:500. doi: 10.3389/fnhum.2016.00500
- Stanislaw, H., and Todorov, N. (1999). Calculation of signal detection theory measures. *Behav. Res. Methods Instrum. Comput.* 31, 137–149. doi: 10.3758/BF03207704
- Tagliabue, M., Zorzi, M., Umiltà, C., and Bassignani, F. (2000). The role of long-term-memory and short-term-memory links in the Simon effect. *J. Exp. Psychol. Hum. Percept. Perform.* 26, 648–670. doi: 10.1037/0096-1523.26.2.648
- Tversky, B. (2011). Visualizing thought. *Top. Cogn. Sci.* 3, 499–535. doi: 10.1111/j.1756-8765.2010.01113.x
- Tzelgov, J., Meyer, J., and Henik, A. (1992). Automatic and intentional processing of numerical information. *J. Exp. Psychol. Learn. Mem. Cogn.* 18, 166–179. doi: 10.1037/0278-7393.18.1.166
- van Dijck, J. P., and Fias, W. (2011). A working memory account for spatial-numerical associations. *Cognition* 119, 114–119. doi: 10.1016/j.cognition.2010.12.013
- Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends Cogn. Sci.* 7, 483–488. doi: 10.1016/j.tics.2003.09.002
- Walsh, V. (2015). “A theory of magnitude: the parts that sum to number,” in *The Oxford Handbook of Numerical Cognition*, eds R. C. Kadosh and A. Dowker (New York, NY: Oxford University Press), 552–565. doi: 10.1093/oxfordhob/9780199642342.013.64
- Wilson, T., and Swets, J. A. (1954). A decision - making theory on visual detection. *Psychol. Rev.* 61, 401–409. doi: 10.1037/h0058700
- Winter, B., Matlock, T., Shaki, S., and Fischer, M. H. (2015). Mental number space in three dimensions. *Neurosci. Biobehav. Rev.* 57, 209–219. doi: 10.1016/j.neubiorev.2015.09.005
- Wühr, P., and Seegelke, C. (2018). Compatibility between physical stimulus size and left-right responses: small is left and large is right. *J. Cogn.* 1:17. doi: 10.5334/joc.19
- Zhang, M., Gao, X., Li, B., Yu, S., Gong, T., Jiang, T., et al. (2016). Spatial representation of ordinal information. *Front. Psychol.* 7:505. doi: 10.3389/fpsyg.2016.00505

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Dural, Burhanoglu, Ekinci, Gürbüz, Akın, Can and Çetinkaya. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



The Use of Local and Global Ordering Strategies in Number Line Estimation in Early Childhood

Jaccoline E. Van 't Noordende^{1,2*}, M. J. M. Volman², Paul P. M. Leseman², Korbinian Moeller^{3,4}, Tanja Dackermann³ and Evelyn H. Kroesbergen⁵

¹ Department of Child Development and Education, University of Amsterdam, Amsterdam, Netherlands, ² Department of Special Education: Cognitive and Motor Disabilities, Utrecht University, Utrecht, Netherlands, ³ Leibniz-Institut für Wissensmedien, Tübingen, Germany, ⁴ Department of Psychology, Universität Tübingen, Tübingen, Germany, ⁵ Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, Netherlands

OPEN ACCESS

Edited by:

Frank Domahs,
Philipps-Universität Marburg,
Germany

Reviewed by:

Dale J. Cohen,
University of North Carolina
at Wilmington, United States
Alex M. Moore,
Franklin & Marshall College,
United States

*Correspondence:

Jaccoline E. Van 't Noordende
j.e.vantnoordende@uva.nl

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 22 December 2017

Accepted: 06 August 2018

Published: 18 September 2018

Citation:

Van 't Noordende JE, Volman MJM, Leseman PPM, Moeller K, Dackermann T and Kroesbergen EH (2018) The Use of Local and Global Ordering Strategies in Number Line Estimation in Early Childhood. *Front. Psychol.* 9:1562. doi: 10.3389/fpsyg.2018.01562

A lot of research has been devoted to number line estimation in primary school. However, less is known about the early onset of number line estimation before children enter formal education. We propose that ordering strategies are building blocks of number line estimation in early childhood. In a longitudinal study, children completed a non-symbolic number line estimation task at age 3.5 and 5 years. Two ordering strategies were identified based on the children's estimation patterns: local and global ordering. Local ordering refers to the correct ordering of successive quantities, whereas global ordering refers to the correct ordering of all quantities across the number line. Results indicated a developmental trend for both strategies. The percentage of children applying local and global ordering strategies increased steeply from 3.5 to 5 years of age. Moreover, children used more advanced local and global ordering strategies at 5 years of age. Importantly, level of strategy use was related to more traditional number line estimation outcome measures, such as estimation accuracy and regression fit scores. These results provide evidence that children use dynamic ordering strategies when solving the number line estimation task in early stages of numerical development.

Keywords: numerical development, number line estimation, strategy use, local ordering, global ordering

INTRODUCTION

The oldest known illustration of a number line was published in 1685 in John Wallis' book "Treatise of Algebra." The concept of the number line was an unconventional idea in the 17th century (Núñez, 2011). Its use increased over time and number lines are nowadays commonly used in research and practice. With its increased use, there has also been an increase in theoretical models and analysis methods to evaluate performance on number line tasks. Most of these models were tested in primary school children and adults. However, numerical skills develop even before formal schooling starts (see Raghubar and Barnes, 2017, for a review on the development of early numerical skills). Nevertheless, our knowledge about the development of number line estimation at preschool age is still rather patchy. However, understanding the processes in the early stages is necessary to identify building blocks of later number line estimation performance. Children below 5 years of age are usually not able to estimate the position of symbolic Arabic numbers on a number line, because they do not yet know these numbers, but they may be able to estimate the position of numerosities in a so-called non-symbolic number line estimation task (cf. Kolkman et al., 2013).

In the current study, 3.5- to 5-year-old children's performance in non-symbolic number line estimation was evaluated with the aim of identifying building blocks of later number line estimation performance. We propose a new method to evaluate children's estimation patterns based on ordering strategies.

The Number Line Estimation Task

In number line estimation tasks, children usually have to estimate the spatial position of numbers on an otherwise empty number line. This number line is usually marked with a numerical start- and end-point (e.g., 0 and 100), although there are also unbounded versions of the number line task (e.g., Cohen and Blanc-Goldhammer, 2011; Cohen and Sarnecka, 2014; Link et al., 2014; Reinert et al., 2015; Opfer et al., 2016). Traditionally, symbolic numbers (i.e., Arabic numerals) are used in number line estimation tasks, but recently non-symbolic quantities (e.g., sets of dots) were used as well (e.g., Kolkman et al., 2013; Fazio et al., 2014; Friso-van den Bos et al., 2014b; Sasanguie et al., 2016). The current study focuses on bounded non-symbolic number line estimation.

Using non-symbolic quantities provides the opportunity to use number line estimation tasks in young children who do not yet master symbolic (Arabic) numbers. Nevertheless, the development of non-symbolic number line estimation has only been studied in children from age 5 years onward (Sasanguie et al., 2012a,b, 2013; Praet and Desoete, 2014; Sella et al., 2015). Most of these studies are based on research on symbolic number line estimation as regards to theoretical background, but also analysis methods are generalized from symbolic to non-symbolic number line estimation. One of the outcome measures used for both symbolic and non-symbolic number line estimation tasks is estimation accuracy, typically operationalized by the percentage absolute error of estimation. This score represents the deviation between participants' estimates and the spatially correct position of the target numbers on the number line. Estimation accuracy was found to increase with age on both symbolic and non-symbolic number line tasks (Siegler and Booth, 2004; Berteletti et al., 2010; Reeve et al., 2015). However, there is an ongoing debate on the underlying cognitive mechanisms that lead to this increase in estimation accuracy. There are two main theoretical accounts: the "mental number line" and the "proportional reasoning" account.

Theoretical Accounts of Number Line Estimation

The mental number line account states that number line estimation performance reflects the underlying mental representations of number magnitude (Siegler and Booth, 2004). This account is based on suggestions by Dehaene (1997, 2001), who states that the basis of numerical cognition is an innate representation of number magnitude in the form of a mental number line. This mental number line was suggested to be logarithmically compressed in those without experience with numbers or education (e.g., Pica et al., 2004), resulting in a characteristic estimation pattern: lower numbers are placed farther away from each other than larger numbers and

thus, placement of the numbers on the number line becomes more dense with increasing numbers. According to this view, such a representation of number magnitude is reflected in a logarithmic distribution of young children's estimates in number line estimation (Siegler and Booth, 2004). Such a logarithmic estimation pattern has been found in both symbolic and non-symbolic number line estimation tasks in 5- to 7-year-old children (Praet and Desoete, 2014). Through experience and education children learn that numbers are equidistant, which means that the distance between two adjacent numbers is always the same (e.g., the distance between 1 and 2 is equal to the distance between 91 and 92). Accordingly, this results in a linear distribution of children's estimates along the number line (e.g., Siegler and Booth, 2004). The time point of the shift from a logarithmic to a linear estimation pattern was observed to depend on the number line format and the number range assessed. Praet and Desoete (2014) showed that second graders' estimation patterns on a number line estimation task using (symbolic) Arabic digits fitted best to a linear model, whereas estimation patterns of the same children on a number line estimation task using (non-symbolic) dot patterns fitted best to a logarithmic model. This suggests that the shift from a logarithmic to a linear estimation pattern will take place earlier for (symbolic) Arabic numbers than for (non-symbolic) dot patterns. Other studies showed that the shift also takes place earlier for smaller than for larger number ranges. For example, in the study of Siegler and Opfer (2003), the best fitting model on second and fourth graders' estimates was linear on a 0–100 number line task but logarithmic on a 0–1000 number line task. Therefore, Siegler and Opfer (2003) proposed that multiple mental number representations may coexist at the same time.

The existence of multiple estimation patterns within an age group was confirmed by the study of Bouwmeester and Verkoeijen (2012). However, they did not find a developmental trend from logarithmic to linear estimation patterns from the age of 5–8 years. Some younger children showed quite accurate (linear) estimation patterns on a symbolic 0–100 number line task, whereas some older children showed inaccurate estimation patterns. Moreover, although Bouwmeester and Verkoeijen (2012) did find a group of children showing estimation patterns resembling a logarithmic distribution, estimation patterns fitted better to a cubic model than to a logarithmic model. Estimation patterns of this group of children showed accurate estimates for numbers close to the beginning, midpoint, and endpoint of the number line, which suggests use of proportional reasoning.

The proportional reasoning account argues that participants' number line estimation performance is not a direct reflection of their mental representation of number magnitude. Instead, it claims that number line estimation performance is influenced by proportional reasoning strategies used to solve the task (Cohen and Blanc-Goldhammer, 2011; Bouwmeester and Verkoeijen, 2012; Cicchini et al., 2014; Huber et al., 2014; Hurst et al., 2014). This account implies that participants use reference points (e.g., the middle of the line reflecting the position of 50 for a number line ranging from 0 to 100) to guide their estimates, which has been tested by applying power models to number line estimation data (Barth and Paladino, 2011;

Cohen and Blanc-Goldhammer, 2011; Slusser et al., 2013; Rouder and Geary, 2014). For example, Rouder and Geary (2014) found that first graders' estimates on a 0–100 number line task were fitted best by a one-cyclic power model reflecting the use of the beginning and endpoint of the line as reference points for estimation. Second graders' consideration of the midpoint of the line as an additional reference point was reflected by a two-cycle power model (Rouder and Geary, 2014). Contrary to the mental number line account, the proportional reasoning account thus assumes that the estimates are actually formed *during* the task, and can even be influenced by specific task characteristics like the presence of external benchmarks on the number line (Peeters et al., 2017a,b).

The mental number line and the proportional reasoning account were tested against each other in research on symbolic number line estimation. Cyclic power models usually provided a better fit to number line estimation patterns than linear and logarithmic regression models from first or second grade and onward (Barth and Paladino, 2011; Slusser et al., 2013; Rouder and Geary, 2014; Friso-van den Bos et al., 2015). Logarithmic and linear estimation patterns seem to be caused by task characteristics (like the use of a bounded number line) instead of underlying mental representations, and the proportional reasoning account could provide an alternative explanation to seemingly logarithmic and linear estimation patterns (Cohen and Sarnecka, 2014; Cohen and Quinlan, 2018). For example, the developmental shift from logarithmic to linear estimation patterns could be explained by development in using proportional reasoning strategies (e.g., from the use of only the beginning and endpoint of the line as reference points, toward additional use of the midpoint of the line as a reference point), instead of development in underlying mental representations (Cohen and Sarnecka, 2014). Nevertheless, Dackermann et al. (2015) argued that neither one account nor the other may be sufficient in itself to fully explain children's performance in number line estimation. Instead, they propose that number line estimation performance builds on both number magnitude representations and proportional reasoning. Moreover, they argue that familiarity with and understanding the characteristics of numbers is also essential to number line estimation. Several studies showed that numerical familiarity and understanding can even be a valid alternative explanation of seemingly logarithmic estimation patterns (e.g., Ebersbach et al., 2008; Stapel et al., 2015). For example, Ebersbach et al. (2008) demonstrated a link between children's counting range (i.e., the range of numbers children could count correctly) and their estimates in number line estimation. Children were able to estimate numbers correctly on the number line as long as the numbers fell within their counting range. It seems reasonable to assume that it is not the mere knowing of the number words and their sequence that enhances number line estimation performance, but the understanding of the numerical magnitudes of the respective numbers. As such, it might be a combination of children's understanding of ordinality and cardinality of numbers that is important to number line estimation. In this context, ordinality refers to understanding the position of numbers in relation to other numbers, whereas cardinality refers to understanding the

actual magnitude of numbers. As indicated above, understanding both ordinality and cardinality are supposed to corroborate accurate estimations on a number line.

Ordering Strategies in Number Line Estimation

The role of ordinality and cardinality in young children's number line estimation performance was investigated by Sullivan and Barner (2014). They showed that kindergartners already understand the ordinal relation between numbers, even before they are able to make correct cardinal estimates on a symbolic number line ranging from 0 to 100. In particular, Sullivan and Barner (2014) examined whether children estimated each number in relation to the preceding number (i.e., the number that was presented directly before the current number). For example, a child first estimated the target number 30 to be located at the position of about 50 on the number line. Next, the target number 40 had to be estimated. In case the child already understands the ordinal relation between numbers 30 and 40, she/he should be able to estimate the location of 40 more rightward on the number line (i.e., somewhere between 50 and 100), even though this would not be the correct cardinal position relative to the beginning and endpoint of the number line. Sullivan and Barner (2014) found that 5-year-old children produced correct ordinal responses on about 70% of the trials, 6-year-olds on 84% of the trials and 7-year-olds on 93% of the trials, regardless whether they placed the target numbers at the correct cardinal position on the number line. Five-year-olds did make these correct ordinal responses without taking into account the correct relative distance between numbers (how far the number is positioned to the right or left of the preceding number), whereas 6- and 7-year-olds did consider relative distance between numbers. Moreover, many of the children did not only make correct ordinal responses in relation to the directly preceding number, but to almost *all* previously estimated numbers. This indicates that children do not only use trial-by-trial ordering, but also monitor their global ordering of numbers across the line on symbolic number line estimation.

A recent study by Cicchini et al. (2014) confirmed that trial-by-trial ordered responses were observed for non-symbolic number line estimation in 8- to 11-year-old children and adults as well. However, Cicchini et al. (2014) did not assess global ordering. Therefore, it is not yet clear whether both local and global ordering are used in non-symbolic number line estimation as well. Moreover, so far studies evaluating local and global ordering in symbolic and non-symbolic number line estimation only investigated children from 5 years of age and adults. The current study will be the first to examine whether children already use either/or both local (trial-by-trial) and global ordering strategies on non-symbolic number line estimation before they enter primary school.

The Current Study

The aim of the current study was to evaluate the early onset and development of strategy use in number line estimation. Therefore, we evaluated estimation performance on a non-symbolic number line estimation task longitudinally in children

from 3.5 to 5 years of age. In particular, we explored a new method of analyzing children's estimation patterns, based on local and global ordering strategies. Local ordering refers to strategies considering response to preceding trials as reference points whereas global ordering refers to strategies reflecting an increasingly left-to-right ordering of increasing quantities across the number line (cf. Sullivan and Barner, 2014). Similar to the proceeding of Sullivan and Barner (2014), we only focused on (correct) ordering of quantities when coding strategy use, and not on correct cardinal positions on the number line.

In line with Sullivan and Barner (2014), we expected a developmental trend for local ordering strategies, from using only ordinal information (whether the target number should be placed to the right or left of the preceding target number) toward taking into account relative distance between quantities (how far to the right or left of the preceding target number). Additionally, we hypothesized a developmental trend for global ordering strategies as well. In the end, all participants should be able to correctly order all estimates across the number line (cf. Sullivan and Barner, 2014), but young children may not yet able to do this. Nevertheless, we hypothesized that young children should be able to use a basic level of global ordering, when ordering small quantities without differentiating between larger quantities (cf. Moeller and Nuerk, 2011).

To be able to correctly position quantities on the number line, both local and global ordering are probably needed. Therefore, we expected that the development in local and global ordering should be associated. Furthermore, increasing levels of both local and global ordering strategy use should lead to more accurate estimations of quantities on the number line. We used this hypothesis to test the validity of local and global ordering strategies as indications of number line estimation performance, by relating strategy use to more traditional measures of number line estimation such as absolute estimation error and regression fit scores.

MATERIALS AND METHODS

Procedure

The current study was part of a larger longitudinal study¹, consisting of two cohorts, followed from age 7 months to 3.5 years and from age 2.5 to 5 years, respectively. Data collected at age 3.5 and 5 years was considered in the current study. This enabled us to evaluate early onset and development of children's strategies in number line estimation, just before children entered kindergarten at the age of 4 years, and follow this development into kindergarten.

Participants were recruited through the local government. The local government provided addresses of all parents with children in the eligible age range. An invitation letter was sent to all of these parents. Additionally, a small number of parents were recruited through Internet forums on parenting or via friends and family. For each cohort, 60 children with

no indications of physical or mental health problems and born on-term (≥ 37 weeks of gestation) were selected to participate. Participants were selected based on order of application. Written informed consent was obtained from the parents of all children participating in the study. The study was approved by the local ethical research committee.

Data collection took place at our lab by trained master's students following a fixed protocol. Parents were allowed to be present during the entire session, but they were instructed not to give any help to the child to complete the tasks.

Participants

Forty-eight children from cohort 1 and 52 children from cohort 2 participated at age 3.5 years. Data from both cohorts were pooled for the current study, which resulted in a total sample of 100 children. Three children did not complete the number line estimation task and were therefore excluded from analyses. The remaining sample consisted of 63 girls (65%) and 34 boys (35%) at time 1. Their mean age was 3.60 years ($SD = 0.06$ years). Seventy-eight children (80%) were from higher educated families (higher vocational training or university completed).

Forty-five children from cohort 2 were tested again at age 5 years (mean age = 4.94 years, $SD = 0.04$ years). All children attended kindergarten at that time. Two of these children did not have data at 3.5 years (due to the child's non-compliance and due to non-participation because of mother's pregnancy) and were only included in the data analyses at 5 years. Thus, the follow-up sample consisted of 32 girls (71%) and 13 boys (29%). Thirty-nine (87%) children were from higher educated families.

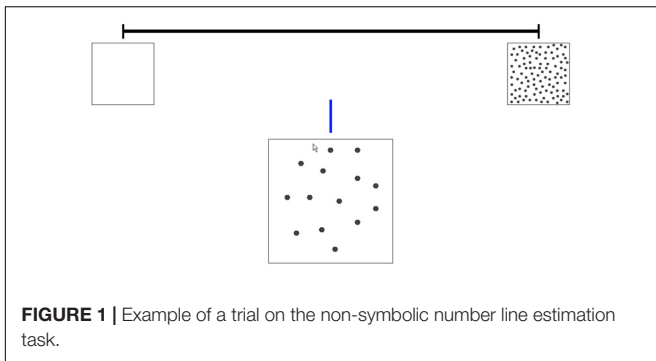
Instruments

An adapted version of the non-symbolic number line task of Kolkman et al. (2013) was used. A line of 1,000 pixels was presented on a computer screen run at a resolution of 1,280 by 1,024 pixels. Only the beginning and endpoint of the number line were marked, with 0 and 100 dots, respectively, throughout the entire task. These quantities were used as a way for children to make sense of the continuum, but were not introduced to the child as the specific numerical quantities "0" and "100." Instead, the experimenter introduced the number line to participants as a road and target quantities as drops of gasoline needed for a car to drive along the road, using terms like "nothing," "a little," "a little more," and "very much."

First, the experimenter presented the child the quantity of 0 and told that the car could not drive without gasoline, and would therefore remain at the startpoint of the road. Next, the experimenter presented a small quantity to the child and pointed out that the car could drive along a small part of the road with this small amount of gasoline. A larger quantity was then presented and the experimenter pointed out that with a larger amount of gasoline the car could drive further along the road. And finally, a quantity of 100 dots was presented and the child was shown that with this large amount of gasoline the car could drive to the end of the road.

Following this instruction, we used four practice trials, in which children had to position quantities, including "0" and

¹ A first draft of this article was published as part of the first author's doctoral thesis (van 't Noordende, 2018).



“100,” upon the number line, to make sure that they understood the concept of the number line.

After practice, participants had to estimate the spatial position of 14 target quantities on the number line. These target quantities (6, 14, 21, 27, 33, 39, 47, 52, 59, 71, 76, 84, 90, and 95) were randomly selected, reflecting an equal distribution across the number range 0–100. The same quantities were used for all participants, but presented in random order. **Figure 1** shows an example of a trial. Quantities to be estimated were presented as dots inside a box below the number line. Dots were equal in size throughout the entire task. As young children might have problems using a computer mouse cursor, participants had to point out the spatial position of each quantity on the number line using his/her finger. The experimenter then dragged the mouse cursor to the position the child indicated.

Analyses

Coding of Strategy Use

Individual estimation patterns were inspected to code the individual level of local and global ordering strategy use. For both strategies, levels were chosen to be mutually exclusive and higher levels were always preferred over lower levels.

Local ordering

To code local ordering strategy use, each estimate was related to the directly preceding estimate to examine whether the ordering of the quantities along the line was correct. Order was considered correct when the estimate was placed correctly to the right or left of the directly preceding estimate on the line. For example, when the first target quantity was 47 and the second target quantity was 33, the second estimate had to be located to the left of the first estimate, regardless whether both estimates were at the correct cardinal position on the line. Note that the target quantities were presented in random order and successive quantities thus differed between children. When the estimate was at about the same position as the previous estimate (i.e., within a 5% range of the number line around the previous estimate's position), it was considered correct in case the numerical difference between the target quantity and the preceding quantity did not exceed 10 (10% of the number line's numerical range). For example, positioning the target quantity 90 and the successive target quantity 95 at the same position was considered correct.

The following levels of local ordering were distinguished (see **Figure 2**):

0. No local ordering
Less than half of the trials were in the correct order compared to the preceding estimate (<7 estimates).
1. Local ordering
More than half of the trials were in the correct order compared to the preceding estimate (≥ 7 estimates).
2. Local ordering with relative distance of $\pm 20\%$
Similar to level 1, but in addition relative distance between correctly ordered successive estimates did not deviate from the correct relative distance between actual target quantity and preceding target quantity by more than $\pm 20\%$ of the number line's numerical range. For example, when the first target quantity was 21 and the second target quantity was 52, the correct relative distance is 31. To fulfill the requirements of this level, the difference between the first and the second estimate has to be between $31 \pm 20 = 11\text{--}51$, regardless whether both estimates are at the correct cardinal position on the number line. In this case, for example, the first estimate could be 26 and the second estimate 46, resulting in a relative distance of 20, which falls in the range of a tolerated relative distance of 11–51.
3. Local ordering with relative distance 10%
Similar to level 2, but the relative distance between correctly ordered successive estimates did not deviate from the correct relative distance between actual target quantity and preceding target quantity by more than $\pm 10\%$ of the number line's numerical range. For example, when the numerical difference between the target quantity and the preceding target quantity was 20, the difference between the estimated quantity and the preceding estimated quantity had to be between 10 and 30.

Global ordering

In addition to local ordering, estimation patterns of each child were also inspected for the level of global ordering strategy use. A level was assigned when the majority of estimates met the description of the level given below. Four outliers (30%) that did not fit the estimation pattern were allowed, as long as a clear pattern meeting the level's criteria was still visible. The following levels of global ordering were distinguished (see **Figure 3**):

0. No global ordering
Estimates did not show a pattern of global numerical ordering; there was no correct distinction between lower and higher quantities (e.g., all estimates were at about the same position on the line).
1. Global ordering small/large
Smaller quantities and larger quantities were distinguished and grouped together on the number line. The group of larger quantities was positioned to the right of the group of smaller quantities, but the cardinal position of the two groups of estimates was not considered. Identification of groups of smaller and larger quantities was based on visual inspection requiring that two groups of (small and larger) quantities could be clearly distinguished. Therefore, ranges and (cardinal) position of groups of quantities could differ between children.

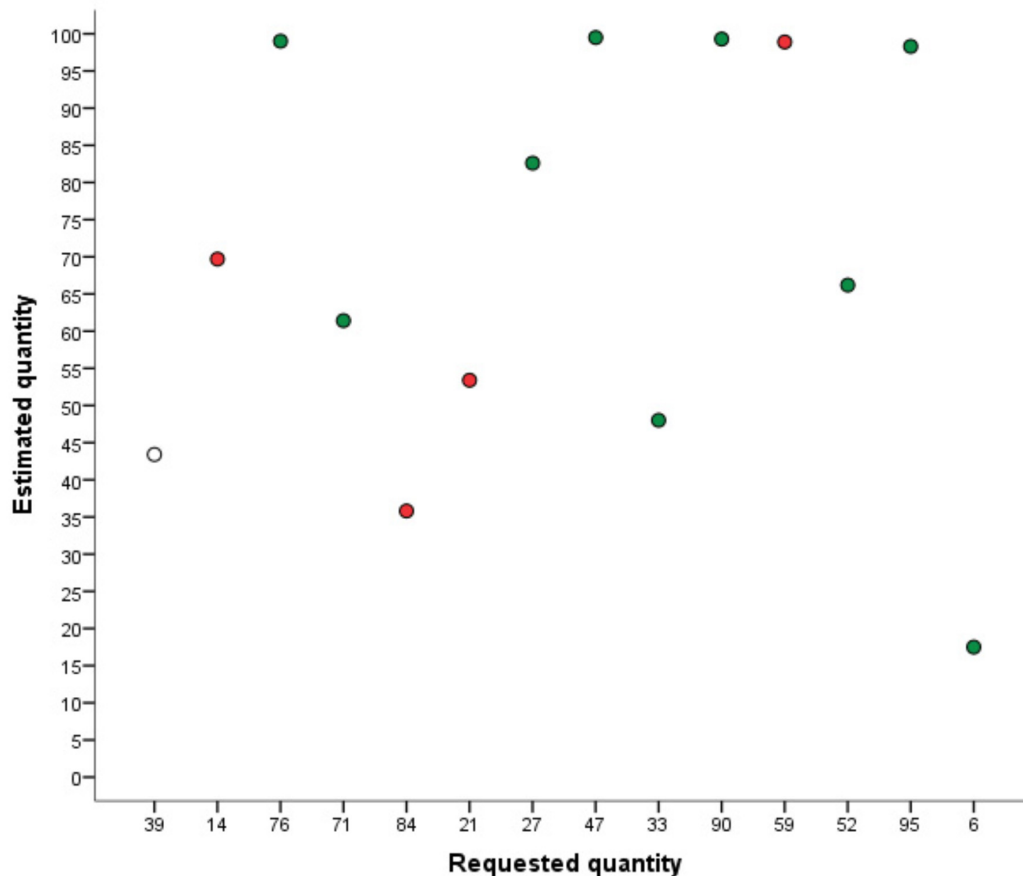


FIGURE 2 | Example of local ordering. The X-axis shows the order of target quantities. Red dots represent incorrectly ordered estimates compared to the directly preceding estimate. Green dots represent correctly ordered estimates.

2. Global ordering small/medium/large

Similar to level 1, but quantities were grouped in three groups from left to right on the number line differentiating small, medium, and large quantities.

3. Global ordering small quantities

Smaller quantities were ordered consecutively, whereas larger quantities were grouped together and not differentiated any further. Larger quantities were positioned to the right of smaller quantities, but cardinal position of estimates was not considered.

4. Global ordering all quantities

The whole range of quantities was ordered consecutively, with larger quantities placed to the right of smaller quantities. Cardinal position of estimates was not considered.

Statistical Analyses

Strategy Use and Development

After coding individual estimation patterns, the results were first analyzed for both time points separately. A frequency distribution indicated the number of children that used the respective strategy levels. Because we hypothesized that local and global ordering

strategies *together* are building blocks of number line estimation, the interrelation between the two strategies was also evaluated, using Kendall's tau-b correlation coefficient².

Next, the development in strategy use from 3.5 to 5 years was investigated for both strategies separately. Wilcoxon signed rank tests were used to evaluate whether the level of local and global ordering strategy use was higher at 5 years than at 3.5 years of age. Kendall's tau-b correlation coefficients were used to evaluate whether level of strategy use at 3.5 years correlated with level of strategy use at 5 years of age.

Finally, the interrelated development of local and global ordering strategies was evaluated. Kendall's tau-b correlation coefficient was used to evaluate the association of the development in local ordering strategies from age 3.5 to 5 years with the development in global ordering strategies. Development in strategy use was indicated by a difference score subtracting the level of strategy use at 3.5 years from the level of strategy use at 5 years of age. Next, for each time point, each

²We choose to use Kendall's tau-b in all analyses considering non-parametric correlations, because it is usually preferred over Spearman's non-parametric correlation for small data sets and data with a large number of tied ranks (cf. Field, 2013).

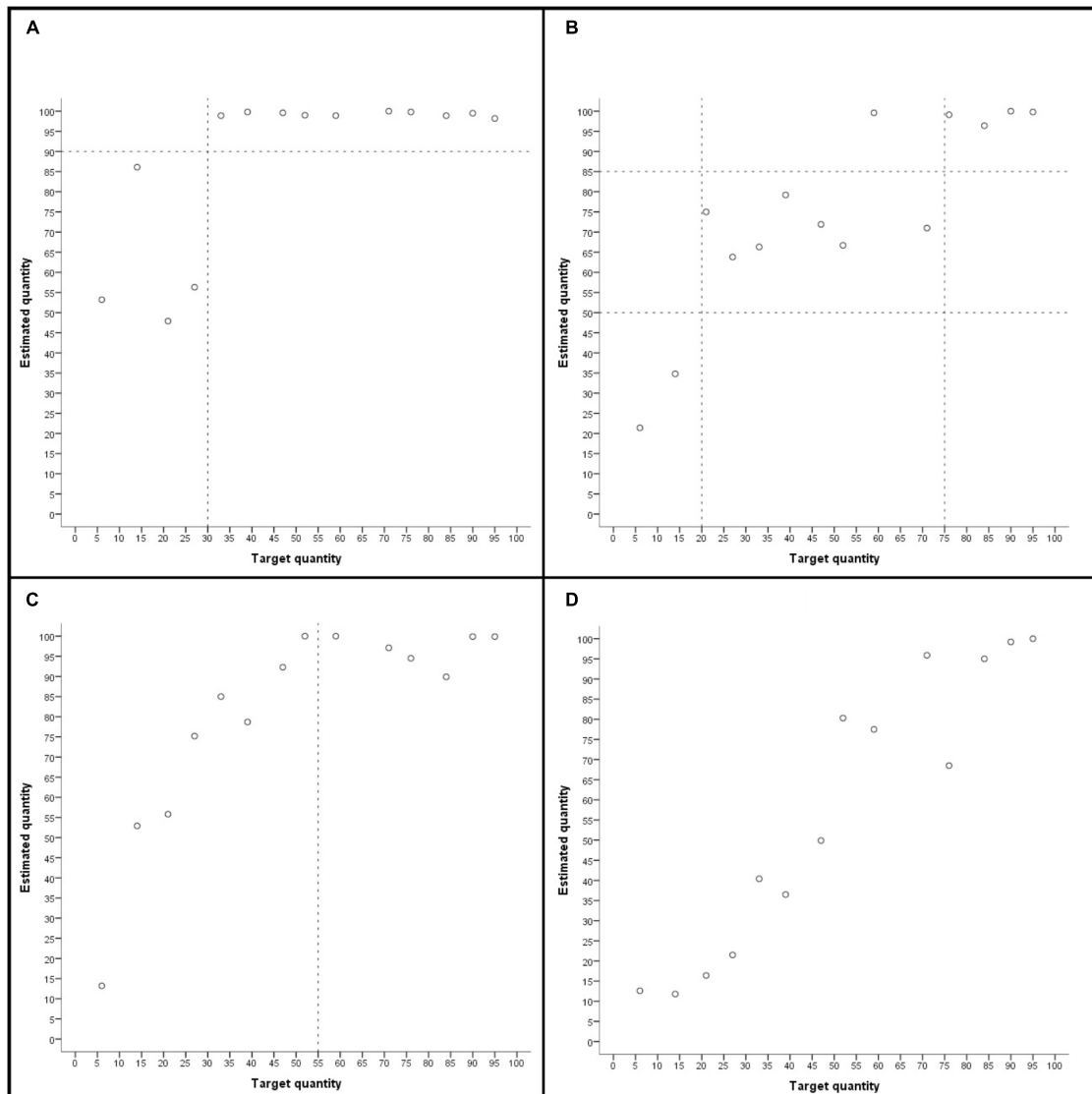


FIGURE 3 | Examples of global ordering level 1 (A), level 2 (B), level 3 (C), and level 4 (D). Dashed lines indicate the distinction between small, medium, and large quantities.

possible combination of the two strategies was assigned to one of seven groups with increasing competence level, by adding the level of local ordering strategy use to the level of global ordering strategy use (e.g., when a child used local ordering level 1 and global ordering level 2, her/his level of strategy combination would be 3). Kendall's tau-b correlation coefficients were used to evaluate the relation between the combined level of strategy use at 3.5 and 5 years. Wilcoxon signed rank tests were used to evaluate whether the level of combined strategy use was higher at 5 years than at 3.5 years.

Relation of Strategy Use and Other Outcome Measures

We hypothesized that local and global ordering strategies are building blocks of number line estimation performance. Higher

level strategies should thus be associated with better estimation performance as indicated by more traditional outcome measures.

First, absolute estimation error was used as an indicator of estimation accuracy. The absolute estimation error was calculated for each item by subtracting the target quantity from the estimated quantity and taking the outcome's absolute value (e.g., when the target quantity was 59 and the child estimated this quantity at position 43 at the line, the absolute estimation error would be $43 - 59 = 16$). Mean absolute estimation error across all items was calculated for each child separately and used as an outcome measure.

Second, model fit of different regression models was considered an indicator of specific estimation patterns. The estimates of each child were regressed onto a linear and a

logarithmic model³. The linear model is thought to reflect more advanced performance than the logarithmic model (as outlined in the introduction). Thus, we expected higher level local and global ordering strategies to be associated with better fit indices of the linear model. Nevertheless, the logarithmic model was tested as well, because children in the current study might be too young to show linear estimation patterns. The individual model fit index R^2 was used as an outcome measure for both models. We would like to emphasize that we do not want to imply an innate mental number line by testing linear and logarithmic regression models. We used these models only as an index of specific data patterns.

Third, the ordinal relation between the target and estimated quantities was quantified using Kendall's tau-b. This non-parametric correlation was used as an alternative to the linear and logarithmic model, to simply evaluate the ordering of estimates without imposing a pre-specified model onto the data.

Before analyzing the relation between strategy use and other outcome measures, absolute estimation error, linear fit index, logarithmic fit index, and ordinal relation index were examined separately. Thirty-one (32.0%) 3.5-year-old children and two (4.4%) 5-year-old children showed a negative correlation between target and estimated quantities. Because all negative relations between the target and estimated quantities are considered incorrect estimation patterns, these children were assigned a score of 0 on the linear fit index, logarithmic fit index, and ordinal relation index. Linear fit index, logarithmic fit index, and ordinal relation index were all heavily skewed to the right at 3.5 years⁴. Therefore, Wilcoxon signed ranks tests and Kendall's tau-b correlation coefficients were used to analyze growth and relation over time of these outcome measures. A dependent samples *t*-test and a Pearson correlation were used to analyze the development of the error score, which was normally distributed at both time points.

The relation between the level of strategy use and the other outcome measures was analyzed using Kendall's tau-b correlation coefficient.

An α -level of 0.05 was used for all statistical analyses.

RESULTS

Strategy Use at 3.5 Years

The frequency distribution of local and global ordering strategy use is depicted in **Table 1**. More than half of the children did not use either a local nor global ordering strategy. For the local ordering strategy, almost all of the remaining children used a local ordering strategy without considering relative distance between quantities (level 1). The variation in level of global ordering

strategy use was larger, but the number of children that used each level of the global ordering strategy decreased from level 1 to level 4.

There was a positive relation between the level of local and global ordering strategy use: $\tau = 0.54$, $p < 0.001$. Most children with a lower level local ordering strategy also used a lower level global ordering strategy. Similarly, children who used a higher level local ordering strategy also used a higher level global ordering strategy. It should be noted, however, that the occurrence of the highest level was quite seldom for global ordering strategies and no child used the highest level of the local ordering strategies.

Strategy Use at 5 Years

Table 2 shows the frequency of strategy use on the non-symbolic number line at 5 years of age. For local ordering strategies, almost all children used either local ordering without relative distance (level 1) or local ordering with 20% relative distance strategy (level 2). Again, there was more variation in levels of global ordering strategies. Frequency of strategy use was quite similar across all levels of global ordering strategy use, although there was a slight increase in frequency from level 1 to level 2 and a slight decrease from level 2 to level 4.

Similar to the results at 3.5 years, Kendall's tau-b showed that levels of local and global ordering strategy use were positively related: $\tau = 0.53$, $p < 0.001$.

Development in Strategy Use

The development in strategy use was first analyzed for the two strategies separately. In general, higher strategies were used at 5 years than at 3.5 years. Only 13% of the 5-year-old children used none of the strategies, compared to 54% of the 3.5-year-old children. A Wilcoxon signed ranks test showed that there was significant improvement in local ordering strategies from 3.5 to 5 years, $z = -4.36$, $p < 0.001$. Twenty-seven children (63%) used a higher local ordering strategy level at 5 years than at 3.5 years of age, as opposed to 13 children (30%) who used the same strategy level at both time points and three children (7%) who used a lower strategy level at 5 years than at 3.5 years of age (see **Table 3**). Nevertheless, there was no significant relation between children's strategy use at 3.5 years and their strategy use at 5 years, as indicated by Kendall's tau-b: $\tau = 0.07$, $p = 0.606$.

The results for global ordering strategy use were similar to the results of local ordering strategy use. Slightly more than half of the children (58%) used a higher global ordering strategy level at 5 years than at 3.5 years of age. Eleven children (26%) used the same strategy at age 3.5 and 5 years of age and seven children (16%) used a lower strategy level at 5 years than at 3.5 years (see **Table 4**). A Wilcoxon signed ranks test showed that the improvement in global ordering strategy level was significant, $z = -3.45$, $p = 0.001$. Again, there was no significant relation between levels of strategy use at 3.5 and 5 years: $\tau = 0.15$, $p = 0.247$.

Next, the interrelation of the development of local and global ordering strategies was investigated by (1) correlating the improvement in local ordering strategy use to the improvement in global ordering strategy use, (2) correlating the combined

³Power models were also tested (cf. Barth and Paladino, 2011; Rouder and Geary, 2014), but one-cycle and two-cycle power models could not be identified for most children. The results of non-cyclic power model were largely identical to those of the logarithmic model. Therefore, power models were not considered here.

⁴The skewed distribution of the linear and logarithmic R^2 was not caused by the scores of 0 that were assigned to children with a negative relation between the target and estimated quantities. The distribution did not change significantly when using the original scores or when excluding these scores. The distribution of Kendall's tau-b was altered by assigning the scores of 0, but the results of the analyses did not change.

TABLE 1 | Frequency of strategy use on the non-symbolic number line at 3.5 years.

	0. No global ordering	1. Global ordering small/large	2. Global ordering small/medium/large	3. Global ordering small quantities	4. Global ordering all quantities	Total
0. No local ordering	52 (53.61)	7 (7.22)	1 (1.03)	0 (0.00)	0 (0.00)	60 (61.86)
1. Local ordering	14 (14.43)	6 (6.19)	6 (6.19)	5 (5.15)	3 (3.09)	34 (35.05)
2. Local ordering relative distance 20%	0 (0.00)	0 (0.00)	1 (1.03)	1 (1.03)	1 (1.03)	3 (3.09)
3. Local ordering relative distance 10%	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)
Total	66 (68.04)	13 (13.40)	8 (8.25)	6 (6.19)	4 (4.12)	97 (100)

TABLE 2 | Frequency of strategy use on the non-symbolic number line at 5 years.

	0. No global ordering	1. Global ordering small/large	2. Global ordering small/medium/large	3. Global ordering small quantities	4. Global ordering all quantities	Total
0. No local ordering	6 (13.33)	2 (4.44)	0 (0.00)	0 (0.00)	0 (0.00)	8 (17.78)
1. Local ordering	3 (6.67)	3 (6.67)	7 (15.56)	4 (8.89)	1 (2.22)	18 (40.00)
2. Local ordering relative distance 20%	0 (0.00)	3 (6.67)	6 (13.33)	3 (6.67)	4 (8.89)	16 (35.56)
3. Local ordering relative distance 10%	0 (0.00)	0 (0.00)	0 (0.00)	3 (6.67)	0 (0.00)	3 (6.67)
Total	9 (20.00)	8 (17.78)	13 (28.89)	10 (22.22)	5 (11.11)	45 (100)

TABLE 3 | Development of local ordering strategy use on the non-symbolic number line from 3.5 to 5 years.

3.5 Years	5 Years				Total
	0. No local ordering	1. Local ordering	2. Local ordering relative distance 20%	3. Local ordering relative distance 10%	
0. No local ordering	5 (11.63)	9 (20.93)	10 (23.26)	1 (2.33)	25 (58.14)
1. Local ordering	2 (4.65)	7 (16.28)	5 (11.63)	2 (4.65)	16 (37.21)
2. Local ordering relative distance 20%	0 (0.00)	1 (2.33)	1 (2.33)	0 (0.00)	2 (4.65)
3. Local ordering relative distance 10%	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)
Total	7 (16.28)	17 (39.53)	16 (37.21)	3 (6.98)	43 (100)

TABLE 4 | Development of global ordering strategy use on the non-symbolic number line from 3.5 to 5 years.

3.5 Years	5 Years					Total
	0. No global ordering	1. Global ordering small/large	2. Global ordering small/medium/large	3. Global ordering small numbers	4. Global ordering all numbers	
0. No global ordering	6 (13.95)	5 (11.63)	7 (16.28)	4 (9.30)	3 (6.98)	25 (58.14)
1. Global ordering small/large	1 (2.33)	2 (4.65)	1 (2.33)	0 (0.00)	2 (4.65)	6 (13.95)
2. Global ordering small/medium/large	1 (2.33)	0 (0.00)	1 (2.33)	3 (6.98)	0 (0.00)	5 (11.63)
3. Global ordering small numbers	0 (0.00)	0 (0.00)	3 (6.98)	2 (4.65)	0 (0.00)	5 (11.63)
4. Global ordering all numbers	0 (0.00)	1 (2.33)	0 (0.00)	1 (2.33)	0 (0.00)	2 (4.65)
Total	8 (18.60)	8 (18.60)	12 (27.91)	10 (23.26)	5 (11.63)	43 (100)

Numbers in parentheses indicate percentage of children of the overall sample.

level of local and global ordering strategy use at 3.5 years with the combined level at 5 years of age, and (3) analyzing the improvement in the combined level of local and global ordering strategy use (see the description of analyses in the Section “Materials and Methods”).

The improvement in local ordering strategies from age 3.5 to 5 years was significantly correlated to improvement in global ordering strategies: $\tau = 0.49$, $p < 0.001$. However, the combined level of strategy use at 3.5 years was not significantly related to the combined level of strategy use at 5 years: $\tau = 0.12$,

$p = 0.325$. Nevertheless, a Wilcoxon signed ranks test showed that there was significant improvement in the combined level of strategy use from 3.5 to 5 years: $z = -3.29$, $p = 0.001$. Twenty-nine children (67%) used a higher level combination of strategies at 5 years compared to 3.5 years. Six children (14%) used the same level combination and eight children (19%) used a lower level combination at 5 years than at 3.5 years.

Relation With Other Outcome Measures

Finally, the relation between strategy use and other (more traditional) outcome measures of the number line estimation task was evaluated. Descriptive statistics and pairwise comparisons for the traditional outcome measures (absolute estimation error, fit indices for linear and logarithmic models, and ordinal relation index) are displayed in **Table 5**. Because linear and logarithmic fit indices as well as the ordinal relation index at 3.5 years were skewed to the right, median and interquartile range are reported for these variables as well. There was significant improvement in performance on all outcome measures (see **Table 5**). Moreover, there was a significant correlation between the ordinal relation index at 3.5 and 5 years, $\tau = 0.22$, $p = 0.049$. Absolute estimation error and linear and logarithmic fit indexes were not significantly correlated over time ($r_{\text{error}} = 0.01$, $p = 0.966$, $\tau_{\text{linear}} = 0.16$, $p = 0.147$, $\tau_{\text{logarithmic}} = 0.17$, $p = 0.121$).

Kendall's tau-b was used to analyze the relation between strategy use and the other outcome measures. Overall, strategy use at 3.5 and 5 years was significantly related to the other outcome measures (see **Table 6**). Use of higher strategy levels was associated with better performance on the other outcome measures. Moreover, the development in strategy use was positively correlated to the development in the other outcome measures.

DISCUSSION

In the current study, we extended previous research on the use of local and global ordering strategies in number line estimation by pursuing an in-depth analysis of the development of non-symbolic number line estimation in 3.5- to 5-year-old children. Generally, the results of the current study indicated that about half of the 3.5-year-old children already made use of local and global ordering when estimating non-symbolic quantities on a number line. However, it needs to be considered that the accuracy of their estimations was low with goodness of fit indices of the linear and logarithmic model as well as the ordinal relation index were heavily skewed to the right, with the majority of scores around 0.

This is in line with previous studies on symbolic and non-symbolic number line estimation, revealing that many young children may not yet have developed the underlying skills of number line estimation sufficiently to make valid estimations (Berteletti et al., 2010; Friso-van den Bos et al., 2014a; Praet and Desoete, 2014; Friso-van den Bos et al., 2015). Nevertheless, in the current study, we observed a significant increase in the percentage of children that used local or global ordering strategies from 3.5 to 5 years of age. The percentage of children that used one or both strategies increased from 46% at 3.5 years to 87% at 5 years. Moreover, different levels of local and global ordering strategy use were identified, following a developmental trend from the use of lower level strategies at age 3.5 years to the use of more advanced strategies at age 5 years, which will be discussed in more detail in the following.

Development of Strategy Use

The developmental trend observed in local ordering in the current study substantiated the results of

TABLE 5 | Descriptive statistics and pairwise comparisons of absolute estimation error, linear and logarithmic fit indexes and ordinal relation index on the non-symbolic number line at 3.5 and 5 years.

	3.5 Years		5 Years		Pairwise comparison
	M	SD	M	SD	
Absolute estimation error	32.82	8.65	25.08	8.60	$p < 0.001^c$
Linear fit	0.16 (0.04 ^a)	0.21 (0.32 ^b)	0.41	0.27	$p < 0.001^d$
Logarithmic fit	0.18 (0.04 ^a)	0.23 (0.39 ^b)	0.46	0.27	$p < 0.001^d$
Ordinal relation	0.21 (0.14 ^a)	0.22 (0.36 ^b)	0.45	0.23	$p < 0.001^d$

$N = 97$ at 3.5 years, $N = 45$ at 5 years. ^aMedian. ^bInterquartile range. ^cDependent samples *t*-test. ^dWilcoxon signed rank test.

TABLE 6 | Kendall's tau-b correlation matrix for the non-symbolic number line at 3.5 and 5 years.

	Local ordering			Global ordering			Strategy combination		
	3.5 Years	5 Years	3.5–5 Years	3.5 Years	5 Years	3.5–5 Years	3.5 Years	5 Years	3.5–5 Years
Absolute estimation error	−0.44***	−0.47***	−0.40**	−0.42***	−0.28*	−0.29*	−0.43***	−0.37**	−0.36**
Linear fit	0.49***	0.59***	0.54***	0.64***	0.48***	0.46***	0.62***	0.55***	0.55***
Logarithmic fit	0.49***	0.52***	0.45***	0.66***	0.43***	0.48***	0.63***	0.48***	0.52***
Ordinal relation	0.50***	0.66***	0.53***	0.63***	0.52***	0.47***	0.62***	0.62***	0.55***

$N = 97$ at 3.5 years, $N = 45$ at 5 years. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Sullivan and Barner (2014). At first, young children seem to primarily consider ordinal information to estimate quantities. They seem to decide where to position the target quantity on the number line based on information whether the current quantity is smaller or larger than the preceding quantity. Note that the preceding quantity refers to the quantity that was *presented* directly before the current target quantity and not necessarily the quantity that precedes the current target quantity numerically. Later in development, children then seemed to take into account relative distance between successive quantities. At this stage, they do not only take into account whether the actual target quantity is smaller or larger than the previous item, but also *how much* it is smaller or larger. In the current study, only 3% of the 3.5-year-old children already considered this in their local ordering strategies. Their estimation pattern reflected correct relative distances between successive quantities, within $\pm 20\%$ of the numerical range of the number line. This percentage increased to 36% at age 5 years. Moreover, some (7%) 5-year-olds even made local ordering responses considering correct relative distance between successive quantities within $\pm 10\%$ of the numerical range of the number line. This suggests that there is not only a developmental trend from simple local ordering to local ordering considering relative distance, but also a developmental trend in the degree at which relative distance is considered.

For global ordering, we focused on the ordering of all quantities along the number line, instead of focusing on trial-by-trial ordering. Based on previous research on logarithmic and linear estimation patterns in symbolic number line estimation (e.g., Booth and Siegler, 2006; Friso-van den Bos et al., 2014a), we expected that global ordering should be observed for small quantities first. In other words, early in development children are expected to only order small quantities consecutively, with no or little distinction between larger quantities. This will be followed by global ordering of the whole range of quantities later in development.

The data partially substantiated our expectation of a developmental trend from global ordering of small quantities to global ordering of all quantities. Both levels of global ordering (ordering small quantities and ordering all quantities) were indeed observed, but a clear developmental trend from ordering small quantities to ordering all quantities was not observed. Generally, frequency of these levels of global ordering was low, especially at 3.5 years of age. It turned out that most 3.5-year-old children only distinguished between small and large quantities or between small, medium, and large quantities in global ordering. At age 5 years, more children were able to differentiate between small quantities or even ordered the whole range of quantities consecutively, but it is likely that the broader developmental transition from ordering small quantities to ordering all quantities takes place beyond the age of the children assessed in the current study.

Despite the clear improvement in estimation performance in non-symbolic number line estimation from 3.5 to 5 years, neither local or global ordering strategies nor the more traditional outcome measures at 3.5 years were significantly associated with

scores at 5 years. This might suggest that the non-symbolic number line task may not measure the same skills at 3.5 and 5 years of age. An alternative explanation for the observed low correlations may be that the way children solve non-symbolic number line estimation *changes* over time. All children showed improvement in their number line estimation performance, but their improvement as well as their future performance could not be predicted significantly by their estimation accuracy at 3.5 years. In this context, it is important to note that many of the 3.5-year-old children did not use local or global ordering estimation strategies at all whereas they did at 5 years—but at various levels. Nevertheless, the correlation between the ordinal relation index at 3.5 and 5 years was significant. This seems to indicate that there is some continuity in the degree of ordering quantities along the number line from age 3.5 to 5 years.

The Relation Between Local and Global Ordering Strategies

Importantly, the present results indicated that local and global ordering strategies do not develop in isolation from each other. We observed that levels of local and global ordering strategy use were highly correlated. This means that with increasing level of local ordering strategy children also used a higher level of global ordering and vice versa. This could indicate that *together* local and global ordering act as building blocks of number line estimation performance. However, it is not yet clear whether the association between local and global ordering is caused by developmental processes or is a necessary artifact of the operationalization of the two strategies. It might be possible that global ordering is not possible without local ordering. Nevertheless, the data seems to indicate that local and global ordering do not necessarily need to reflect the same level of proficiency at each time point. Some children used no local ordering but did use global ordering or vice versa. Some children even used one of the lower levels of one strategy and one of the higher levels of the other strategy.

To clarify the issue of dependency of the strategies, we ran some simulations (see **Appendix 1** for the simulation procedure and results). In particular, we simulated local ordering at different levels (i.e., 100 simulated participants for each level of local ordering) and then coded global ordering for these simulated estimation patterns. The results of this simulation were similar to the results observed in our data. The correlation (as indicated by Kendall's tau-b) between simulated local and global ordering strategies was 0.68. Despite this high correlation, the frequency table of simulated strategies showed considerable variation in the level of global ordering strategies within each level of local ordering, except for local ordering level 0. For local ordering level 0, only 6 out of 100 simulated estimation patterns were coded as global ordering. This might lead to the conclusion that various levels of global ordering arise from local ordering by chance. However, both the simulation data and the participants' data also showed that it is difficult, but not impossible, to achieve global ordering without local ordering.

This seems to substantiate that, although local and global ordering are related, they are not just two sides of the same coin.

Furthermore, an important theoretical distinction between the two strategies can be made. Local ordering assumes the use of previous estimates as reference points, whereas global ordering assumes the use of external reference points, like the beginning or endpoint of the number line. For example, in global ordering level 1, small and large quantities are distinguished, with large quantities positioned rightward of small quantities. This requires relating quantities to the beginning and/or endpoint of the line to decide where to position small and large quantities on the number line.

Nevertheless, it is possible that local and global ordering strategies become more integrated over time. The results of the current study showed that children's local ordering strategies developed from ordering successive quantities toward taking into account the relative distance between quantities. Estimating relative distance requires taking into account the length of the number line to estimate the proportion of the line that corresponds to the relative distance between quantities. Therefore, quantities have to be related to external reference points on the number line, like the beginning and endpoint of the line. When the distance between these external reference points is not taken into account it would probably not be possible to estimate the correct proportion of the number line that corresponds to the relative distance between quantities. This resembles first steps toward proportion-based estimation strategies, which have previously been demonstrated to be solution strategies in number line estimation, as indicated by fitting of cyclic power models to estimation patterns (e.g., Barth and Paladino, 2011). The current study extends these previous findings by suggesting that proportion-based estimation strategies may also incorporate previous estimates as reference points.

In the current study, cyclic power models could not be identified reliably. These models probably require more advanced proportional reasoning. So far, reliable fit of cyclic power models was usually found from first grade onward (e.g., Friso-van den Bos et al., 2015). We hypothesize that children will increasingly make use of both previous estimates and external benchmarks on the number line as reference points for estimation throughout development. Together, local and global ordering strategies should act as building blocks of number line estimation. In line with this notion, the current study indicated that higher level local and global ordering was associated with improved estimation performance in non-symbolic number line estimation. Children who used higher levels of local and global ordering also showed higher estimation accuracy, higher logarithmic and linear fit scores, as well as a higher ordinal relation between target and estimated quantities on their non-symbolic number line estimation at 3.5 and 5 years of age. Future research is needed to further investigate the interplay between using previous estimates and external reference points, in order to better understand the relation between local and global ordering, and their role as building blocks of number line estimation performance, throughout development.

Nevertheless, our findings support the view that estimates may be formed during task execution (Cohen and Sarnecka, 2014; Cohen and Quinlan, 2018), and seem to offer an alternative explanation of seemingly logarithmic and linear estimation patterns found in the current study. For example, global ordering of small quantities without differentiating between larger quantities (i.e., global ordering level 3 in the current study) would result in a seemingly logarithmic estimation pattern. Similarly, global ordering of all quantities would result in a seemingly linear estimation pattern. Therefore, even though strategy use was related to logarithmic and linear fit scores in the current study, these estimation patterns may not necessarily reflect mental number line representations, but can instead be explained by strategy use. The current study thus showed that number line estimation does not seem to be a unidimensional construct, but rather builds on interacting strategies, stressing the importance of on-task processing and strategy use instead of mental number line representations.

Underlying Mechanisms of Strategy Use

The important role of dynamic ordering strategies in number line estimation might suggest that children's estimates are primarily guided by ordinal processes at these early ages. Nevertheless, other processes might play a role in local and global ordering strategy use as well. Although the underlying mechanisms of strategy use were not investigated in the current study, we would like to make some suggestions based on previous research to specify potential starting points for further research.

As discussed above, refining estimation of relative position of quantities on a number line probably requires general cognitive skills like analogical and proportional reasoning as well (e.g., Barth and Paladino, 2011; Sullivan and Barner, 2014). Moreover, the use of reference points probably also requires working memory, as for example participants have to remember the position of previous estimates. Therefore, we hypothesize that general cognitive skills play an important role in number line estimation.

Nevertheless, domain-specific numerical skills are also needed to estimate numbers on a number line. It is likely that both ordinality and cardinality are underlying mechanisms in children's non-symbolic number line estimation. To use local ordering, mainly ordinal information is used as participants compare the target quantity to the preceding quantity and decide which one is smaller and which one is larger. For example, when ordering quantities 71 and 75, participants have to understand that the second quantity is larger than the first, but not necessarily that the first quantity is 71 and the second quantity is 75. For global ordering, cardinal processes might play an important role. Results of the current study showed that in general smaller quantities were placed closer to the beginning point and larger quantities were placed closer to the endpoint of the number line. This was already observed at the lowest levels of global ordering. This might indicate that in global ordering children considered not only the relation between quantities, but also the actual magnitudes when considering the relative distance between quantities and external reference points of the number line. This is in line with propositions in previous research on

the use of proportional reasoning strategies in number line estimation (e.g., Barth and Paladino, 2011; Sullivan and Barner, 2014). Even if participants did not estimate the correct cardinal position on the line, this suggests that the cardinal value of each quantity is considered when estimating relative distance between the target quantity and external reference points on the number line. Interestingly, Lyons and Beilock (2013) proposed that non-symbolic ordinal tasks are actually solved through considering cardinality as well. As such, in local ordering non-symbolic quantities may be ordered by comparing the cardinal value of each quantity with the preceding quantity, instead of relating the quantities to their “neighbors.” More research is needed to clarify the role of cardinality and ordinality in non-symbolic number line estimation.

Another skill that is probably needed for number line estimation is visual discrimination of quantities and classification of the difference between quantities in terms of smaller and larger. In case a child cannot discriminate between quantities, she/he would not be able to place quantities on the line in an ordered manner. However, the relation between quantity discrimination and non-symbolic number line estimation is not yet clear. Some studies have shown that non-symbolic quantity discrimination and non-symbolic number line estimation are associated (Kolkman et al., 2013; Friso-van den Bos et al., 2014b), while others have proposed that these tasks rely on different underlying mechanisms (Sasanguie and Reynvoet, 2013). Further research is needed to evaluate the relation between quantity discrimination and non-symbolic number line estimation.

Limitations of the Current Study

When interpreting the results of the current study it is important to note that almost all participants were from rather high SES families. This limits the external validity of the current study to other SES classes, because cognitive performance was found to be influenced by SES (e.g., Jordan et al., 2006). Furthermore, it is not known to what extent children in the current study were able to discriminate between respective quantities. As mentioned above, discrimination of quantities might be related to number line estimation performance. Previous research showed that 3-year-olds are able to discriminate quantities at the ratio of 3:4 and 5-year-olds are able to discriminate quantities at the ratio of 4:5 (Halberda and Feigenson, 2008). However, the results of Halberda and Feigenson (2008) might not be transferable to our items easily, because Halberda and Feigenson (2008) controlled their stimuli for non-numerical cues like surface area, etc., whereas stimuli in the current study were not controlled for these cues. Instead, dot size was kept constant across stimuli, which resulted in a positive association between numerical quantity and total surface area (i.e., larger quantities cover a larger total surface area). In previous research, visual-spatial cues associated with numerical quantity were controlled in non-symbolic stimuli to make children attend to numerical quantity instead of continuous extent (introduced by Clearfield and Mix, 1999). However, the ecological validity of such controlled stimuli might be low. Instead, it is likely that visual-spatial extent and numerical quantity are hard to separate (see Leibovich et al., 2017, for a discussion). Following Cantrell and Smith (2013), we argue

that the association between continuous extent and numerical quantity may not be a problem, but makes numerical quantity more salient to participants on non-symbolic numerical tasks, as an ecologically valid aid. As such, children should have had fewer difficulties discriminating quantities in the current study.

CONCLUSION

Taken together, the current study provides a new perspective on number line estimation in early childhood. The results indicate that the seemingly logarithmic (and linear) patterns found in previous research do not necessarily represent static mental number representations, but may instead be explained by children's dynamic ordering strategies while performing the task. Furthermore, it suggests that the logarithmic estimation pattern often observed for young children and unfamiliar number ranges (e.g., Siegler and Opfer, 2003; Siegler and Booth, 2004) does not seem to be the most basic form of number line estimation. Even before children can order small quantities consecutively, they are able to differentiate between small and large or small, medium and large quantities on the number line. Non-symbolic number line estimation hence builds on the use of local and global ordering strategies, which are already present at 3.5 years of age. These strategies develop from simply considering the ordinality of target quantities to more complex levels of local and global ordering strategies also considering first aspects of cardinality and proportional reasoning between the age of 3.5 and 5 years.

Importantly, we suggest that these strategies represent building blocks, not an end stage of non-symbolic number line estimation. Local and global ordering strategies as measured in the current study may only represent early and basic levels of strategy use. For example, the highest level of global ordering in the current study was assigned when a child ordered all quantities correctly, even when relative distance between quantities or the cardinal position of quantities on the number line was not correct. Considering these aspects would require further development of the respective strategy levels. Future research may therefore aim to incorporate correct ordering as well as correct relative distance between quantities and correct (cardinal) position on the number line, in particular when studying older children. Furthermore, as symbolic number skills become more important during primary school, it would be desirable to also investigate the generalizability of local and global ordering strategies to symbolic number line estimation.

AUTHOR CONTRIBUTIONS

JVN, MV, PL, and EK contributed to the outline of this research. JVN collected the data and supervised the study. All authors contributed to data analysis and writing.

FUNDING

This research was funded by the Netherlands Organisation for Scientific Research (NWO) (404-10-506).

REFERENCES

- Barth, H. C., and Paladino, A. M. (2011). The development of numerical estimation: evidence against a representational shift. *Dev. Sci.* 14, 125–135. doi: 10.1111/j.1467-7687.2010.00962.x
- Berteletti, I., Lucangeli, D., Piazza, M., Dehaene, S., and Zorzi, M. (2010). Numerical estimation in preschoolers. *Dev. Psychol.* 46, 545–551. doi: 10.1037/a0017887
- Booth, J. L., and Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Dev. Psychol.* 42, 189–201. doi: 10.1037/0012-1649.41.6.189
- Bouwmeester, S., and Verkoeijen, P. P. J. L. (2012). Multiple representations in number line estimation: a developmental shift or classes of representations? *Cognit. Instruct.* 30, 246–260. doi: 10.1080/07370008.2012.689384
- Cantrell, L., and Smith, L. B. (2013). Open questions and a proposal: a critical review of the evidence on infant numerical abilities. *Cognition* 128, 331–352. doi: 10.1016/j.cognition.2013.04.008
- Cicchini, G. M., Anobile, G., and Burr, D. C. (2014). Compressive mapping of number to space reflects dynamic encoding mechanisms, not static logarithmic transform. *Proc. Natl. Acad. Sci. U.S.A.* 111, 7867–7872. doi: 10.1073/pnas.1402785111
- Clearfield, M., and Mix, K. (1999). Number versus contour length in infants' discrimination of small visual sets. *Psychol. Sci.* 10, 408–411. doi: 10.1111/1467-9280.00177
- Cohen, D. J., and Blanc-Goldhammer, D. (2011). Numerical bias in bounded and unbounded number line tasks. *Psychon. Bull. Rev.* 18, 331–338. doi: 10.3758/s13423-011-0059-z
- Cohen, D. J., and Quinlan, P. T. (2018). The log-linear response function of the bounded number-line task is unrelated to the psychological representation of quantity. *Psychon. Bull. Rev.* 25, 447–454. doi: 10.3758/s13423-017-1290-z
- Cohen, D. J., and Sarnecka, B. W. (2014). Children's number-line estimation shows development of measurement skills (not number representations). *Dev. Psychol.* 50, 1640–1652. doi: 10.1037/a0035901
- Dackermann, T., Huber, S., Bahnmüller, J., Nuerk, H.-C., and Moeller, K. (2015). An integration of competing accounts on children's number line estimation. *Front. Psychol.* 6:884. doi: 10.3389/fpsyg.2015.00884
- Dehaene, S. (1997). *The Number Sense: How the Mind Creates Mathematics*. New York, NY: Oxford University Press.
- Dehaene, S. (2001). Précis of the number sense. *Mind Lang.* 16, 16–36. doi: 10.1111/1468-0017.00154
- Ebersbach, M., Luwel, K., Frick, A., Onghena, P., and Verschaffel, L. (2008). The relationship between the shape of the mental number line and familiarity with numbers in 5-to 9-year old children: evidence for a segmented linear model. *J. Exp. Child Psychol.* 99, 1–17. doi: 10.1016/j.jecp.2007.08.006
- Fazio, L. K., Bailey, D. H., Thompson, C. A., and Siegler, R. S. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement. *J. Exp. Child Psychol.* 123, 53–72. doi: 10.1016/j.jecp.2014.01.013
- Field, A. (2013). *Discovering Statistics Using IBM SPSS statistics*. London: SAGE.
- Friso-van den Bos, I., Kolkman, M. E., Kroesbergen, E. H., and Leseman, P. P. M. (2014a). Explaining variability: numerical representations in 4- to 8-year-old children. *J. Cognit. Dev.* 15, 325–344. doi: 10.1080/15248372.2012.742900
- Friso-van den Bos, I., Kroesbergen, E. H., and van Luit, J. E. H. (2014b). Number sense in kindergarten children: factor structure and working memory predictors. *Learn. Ind. Diff.* 33, 23–29. doi: 10.1016/j.lindif.2014.05.003
- Friso-van den Bos, I., Kroesbergen, E. H., Van Luit, J. E. H., Xenidou-Dervou, I., Jonkman, L. M., et al. (2015). Longitudinal development of number line estimation and mathematics performance in primary school children. *J. Exp. Child Psychol.* 134, 12–29. doi: 10.1016/j.jecp.2015.02.002
- Halberda, J., and Feigenson, L. (2008). Developmental change in the acuity of the “number sense”: the approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Dev. Psychol.* 44, 1457–1465. doi: 10.1037/a0012682
- Huber, S., Moeller, K., and Nuerk, H. (2014). Dissociating number line estimations from underlying numerical representations. *Quart. J. Exp. Psychol.* 67, 991–1003. doi: 10.1080/17470218.2013.838974
- Hurst, M., Monahan, K. L., Heller, E., and Cordes, S. (2014). 123s and ABCs: developmental shifts in logarithmic-to-linear responding reflect fluency with sequence values. *Dev. Sci.* 17, 892–904. doi: 10.1111/desc.12165
- Jordan, N., Kaplan, D., Olah, L., and Locuniak, M. (2006). Number sense growth in kindergarten: a longitudinal investigation of children at risk for mathematics difficulties. *Child Dev.* 77, 153–175. doi: 10.1111/j.1467-8624.2006.00862.x
- Kolkman, M. E., Kroesbergen, E. H., and Leseman, P. P. M. (2013). Early numerical development and the role of non-symbolic and symbolic skills. *Learn. Instruct.* 25, 95–103. doi: 10.1016/j.learninstruct.2012.12.001
- Leibovich, T., Katzin, N., Harel, M., and Henik, A. (2017). From “sense of number” to “sense of magnitude”: the role of continuous magnitudes in numerical cognition. *Behav. Brain Sci.* 40 1–16. doi: 10.1017/S0140525X16000960
- Link, T., Huber, S., Nuerk, H.-C., and Moeller, K. (2014). Unbounding the mental number line-new evidence on children's spatial representation of numbers. *Front. Psychol.* 4:1021. doi: 10.3389/fpsyg.2013.01021
- Lyons, I. M., and Beilock, S. L. (2013). Ordinality and the nature of symbolic numbers. *J. Neurosci.* 33, 17052–17061. doi: 10.1523/JNEUROSCI.1775-13.2013
- Moeller, K., and Nuerk, H. (2011). Psychophysics of numerical representation: why seemingly logarithmic representations may rather be multi-linear. *Z. Psychol. J. Psychol.* 219, 64–70. doi: 10.1027/2151-2604/a000048
- Núñez, R. E. (2011). No innate number line in the human brain. *J. Cross-Cult. Psychol.* 42, 651–668. doi: 10.1177/0022022111406097
- Opfer, J. E., Thompson, C. A., and Kim, D. (2016). Free versus anchored numerical estimation: a unified approach. *Cognition* 149, 11–17. doi: 10.1016/j.cognition.2015.11.015
- Peeters, D., Sekeris, E., Verschaffel, L., and Luwel, K. (2017a). Evaluating the effect of labeled benchmarks on children's number line estimation performance and strategy use. *Front. Psychol.* 8:1082. doi: 10.3389/fpsyg.2017.01082
- Peeters, D., Verschaffel, L., and Luwel, K. (2017b). Benchmark-based strategies in whole number line estimation. *Br. J. Psychol.* 108, 668–686. doi: 10.1111/bjop.12233
- Pica, P., Lemer, C., Izard, W., and Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science* 306, 499–503. doi: 10.1126/science.1102085
- Praet, M., and Desoete, A. (2014). Number line estimation from kindergarten to grade 2: a longitudinal study. *Learn. Instruct.* 33, 19–28. doi: 10.1016/j.learninstruct.2014.02.003
- Raghubar, K. P., and Barnes, M. A. (2017). Early numeracy skills in preschool-aged children: a review of neurocognitive findings and implications for assessment and intervention. *Clin. Neuropsychol.* 31, 329–351. doi: 10.1080/13854046.2016.1259387
- Reeve, R. A., Paul, J. M., and Butterworth, B. (2015). Longitudinal changes in young children's 0-100 to 0-1000 number-line error signatures. *Front. Psychol.* 6:647. doi: 10.3389/fpsyg.2015.00647
- Reinert, M., Huber, S., Nuerk, H.-C., and Moeller, K. (2015). Strategies in unbounded number line estimation? evidence from eye-tracking. *Cognit. Process.* 16, S359–S363. doi: 10.1007/s10339-015-0675-z
- Rouder, J. N., and Geary, D. C. (2014). Children's cognitive representation of the mathematical number line. *Dev. Sci.* 17, 525–536. doi: 10.1111/desc.12166
- Sasanguie, D., De Smedt, B., Defever, E., and Reynvoet, B. (2012a). Association between basic numerical abilities and mathematics achievement. *Br. J. Dev. Psychol.* 30, 344–357. doi: 10.1111/j.2044-835X.2011.02048.x
- Sasanguie, D., Van den Bussche, E., and Reynvoet, B. (2012b). Predictors for mathematics achievement? evidence from a longitudinal study. *Mind Brain Educ.* 6, 119–128. doi: 10.1111/j.1751-228X.2012.01147.x
- Sasanguie, D., Göbel, S. M., Moll, K., Smets, K., and Reynvoet, B. (2013). Approximate number sense, symbolic number processing, or number-space mappings: what underlies mathematics achievement? *J. Exp. Child Psychol.* 114, 418–431. doi: 10.1016/j.jecp.2012.10.012
- Sasanguie, D., and Reynvoet, B. (2013). Number comparison and number line estimation rely on different mechanisms. *Psychol. Belg.* 53, 17–35. doi: 10.5334/pb-53-4-17
- Sasanguie, D., Verschaffel, L., Reynvoet, B., and Luwel, K. (2016). The development of symbolic and non-symbolic number line estimations: three developmental accounts contrasted within cross-sectional and longitudinal data. *Psychol. Belg.* 56, 382–405. doi: 10.5334/pb.276
- Sella, F., Berteletti, I., Lucangeli, D., and Zorzi, M. (2015). Varieties of quantity estimation in children. *Dev. Psychol.* 51, 758–770. doi: 10.1037/a0039183

- Siegler, R. S., and Booth, J. L. (2004). Development of numerical estimation in young children. *Child Dev.* 75, 428–444. doi: 10.1111/j.1467-8624.2004.00684.x
- Siegler, R. S., and Opfer, J. E. (2003). The development of numerical estimation: evidence for multiple representations of numerical quantity. *Psychol. Sci.* 14, 237–243. doi: 10.1111/1467-9280.02438
- Slusser, E. B., Santiago, R. T., and Barth, H. C. (2013). Developmental change in numerical estimation. *J. Exp. Psychol. Gen.* 142, 193–208. doi: 10.1037/a0028560
- Stapel, J. C., Hunnius, S., Bekkering, H., and Lindemann, O. (2015). The development of numerosity estimation: evidence for a linear number representation early in life. *J. Cognit. Psychol.* 27, 400–412. doi: 10.1080/20445911.2014.995668
- Sullivan, J., and Barner, D. (2014). The development of structural analogy in number-line estimation. *J. Exp. Child Psychol.* 128, 171–189. doi: 10.1016/j.jecp.2014.07.004
- van 't Noordende, J. E. (2018). *Building Blocks of Numerical Cognition: The Development of Quantity-Space Mapping*. Doctoral thesis, Utrecht University, Utrecht. Available at: <https://dspace.library.uu.nl/handle/1874/364782>
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Van 't Noordende, Volman, Leseman, Moeller, Dackermann and Kroesbergen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

APPENDIX 1

The computer program R was used to simulate each level of local ordering, using the following codes, in which the variable x refers to the target quantities and the variable y refers to simulated estimates:

Local ordering level 0:

```
x <- sample(c(6, 14, 21, 27, 33, 39, 47, 52, 59, 71, 76, 84, 90, 95), 14)
```

```
y <- sample(c(0:100), 14)
```

Local ordering level 1:

```
x <- sample(c(6, 14, 21, 27, 33, 39, 47, 52, 59, 71, 76, 84, 90, 95), 14)
```

```
y1 <- sample(0:100, 1)
```

```
if (x[2] < x[1]) y2 <- sample(c((0:(y1-(abs(x[2]-x[1]) + 20))), ((y1-(abs(x[2]-x[1]) - 20)):(y1-5))), 1)
```

```
if (x[2] > x[1]) y2 <- sample(c((y1+5):(y1+(abs(x[2]-x[1]) - 20))), ((y1+(abs(x[2]-x[1]) + 20)):100)), 1)
```

```
if (y2 < 0) y2 <- 0
```

```
if (y2 > 100) y2 <- 100
```

```
if (x[3] < x[2]) y3 <- sample(c((0:(y2-(abs(x[3]-x[2]) + 20))), ((y2-(abs(x[3]-x[2]) - 20)):(y2-5))), 1)
```

```
if (x[3] > x[2]) y3 <- sample(c((y2+5):(y2+(abs(x[3]-x[2]) - 20))), ((y2+(abs(x[3]-x[2]) + 20)):100)), 1)
```

```
if (y3 < 0) y3 <- 0
```

```
if (y3 > 100) y3 <- 100
```

And so on for all 14 items.

Local ordering level 2:

```
x <- sample(c(6, 14, 21, 27, 33, 39, 47, 52, 59, 71, 76, 84, 90, 95), 14)
```

```
y1 <- sample(0:100, 1)
```

```
if (x[2] < x[1]) y2 <- y1 - sample((abs(x[2]-x[1]) - 20):(abs(x[2]-x[1]) + 20), 1)
```

```
if (x[2] > x[1]) y2 <- y1 + sample((abs(x[2]-x[1]) - 20):(abs(x[2]-x[1]) + 20), 1)
```

```
if (y2 < 0) y2 <- 0
```

```
if (y2 > 100) y2 <- 100
```

```
if (x[3] < x[2]) y3 <- y2 - sample((abs(x[3]-x[2]) - 20):(abs(x[3]-x[2]) + 20), 1)
```

```
if (x[3] > x[2]) y3 <- y2 + sample((abs(x[3]-x[2]) - 20):(abs(x[3]-x[2]) + 20), 1)
```

```
if (y3 < 0) y3 <- 0
```

```
if (y3 > 100) y3 <- 100
```

And so on for all 14 items.

Local ordering level 3:

```
x <- sample(c(6, 14, 21, 27, 33, 39, 47, 52, 59, 71, 76, 84, 90, 95), 14)
```

```
y1 <- sample(0:100, 1)
```

```
if (x[2] < x[1]) y2 <- y1 - sample((abs(x[2]-x[1]) - 10):(abs(x[2]-x[1]) + 10), 1)
```

```
if (x[2] > x[1]) y2 <- y1 + sample((abs(x[2]-x[1]) - 10):(abs(x[2]-x[1]) + 10), 1)
```

```
if (y2 < 0) y2 <- 0
```

```
if (y2 > 100) y2 <- 100
```

```
if (x[3] < x[2]) y3 <- y2 - sample((abs(x[3]-x[2]) - 10):(abs(x[3]-x[2]) + 10), 1)
```

```
if (x[3] > x[2]) y3 <- y2 + sample((abs(x[3]-x[2]) - 10):(abs(x[3]-x[2]) + 10), 1)
```

```
if (y3 < 0) y3 <- 0
```

```
if (y3 > 100) y3 <- 100
```

And so on for all 14 items.

TABLE A1 | This resulted in the following frequency table:

	0. No global ordering	1. Global ordering small/large	2. Global ordering small/medium/large	3. Global ordering small quantities	4. Global ordering all quantities	Total
0. No local ordering	94 (23.50)	6 (1.50)	0 (0.00)	0 (0.00)	0 (0.00)	100 (25.00)
1. Local ordering	30 (7.50)	42 (10.50)	26 (6.50)	2 (0.50)	0 (0.00)	100 (25.00)
2. Local ordering relative distance 20%	3 (0.75)	16 (4.00)	59 (14.75)	16 (4.00)	6 (1.50)	100 (25.00)
3. Local ordering relative distance 10%	1 (0.25)	13 (3.25)	56 (14.00)	11 (2.75)	19 (4.75)	100 (25.00)
Total	33 (8.25)	80 (20.00)	130 (32.50)	29 (7.25)	28 (7.00)	400 (100)

Numbers in parentheses indicate percentage of children of the overall sample.



Differential Development of Children's Understanding of the Cardinality of Small Numbers and Zero

Silvia Pixner^{1*}, Verena Dresen¹ and Korbinian Moeller^{2,3}

¹ Institute of Psychology, UMIT – Private University for Health Sciences, Medical Informatics and Technology, Hall in Tirol, Austria, ² Leibniz-Institut für Wissensmedien, Tübingen, Germany, ³ LEAD Graduate School & Research Network and Department of Psychology, University of Tübingen, Tübingen, Germany

OPEN ACCESS

Edited by:

Frank Domahs,
Philipps-Universität Marburg,
Germany

Reviewed by:

James Negen,
Durham University, United Kingdom
Alyssa J. Kersey,
The University of Chicago,
United States
Kristy VanMarle,
University of Missouri, United States

*Correspondence:

Silvia Pixner
silvia.pixner@umit.at

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 07 December 2017

Accepted: 15 August 2018

Published: 25 September 2018

Citation:

Pixner S, Dresen V and Moeller K
(2018) Differential Development
of Children's Understanding of the
Cardinality of Small Numbers
and Zero. *Front. Psychol.* 9:1636.
doi: 10.3389/fpsyg.2018.01636

Counting and the understanding of cardinality are important steps in children's numerical development. Recent studies have indicated that language and visuospatial abilities play an important role in the development of children's cardinal knowledge of small numbers. However, predictors for the knowledge about zero were usually not considered in these studies. Therefore, the present study investigated whether the acquisition of cardinality knowledge on small numbers and the concept of zero share cross-domain and domain-specific numerical predictors. Particular interest was paid to the question whether visuospatial abilities – in addition to language abilities – were associated with children's understanding of small numbers and zero. Accordingly, we assessed kindergarteners aged 4 to 5 years in terms of their understanding of small numbers and zero as well as their visuospatial, general language, counting, Arabic number identification abilities, and their finger number knowledge. We observed significant zero-order correlations of vocabulary, number identification, finger knowledge, and counting abilities with children's knowledge about zero as well as understanding of the cardinality of small numbers. Subsequent regression analyses substantiated the influences of counting abilities on knowledge about zero and the influences of both counting abilities and finger knowledge on children's understanding of the cardinality of small numbers. No significant influences of cross-domain predictors were observed. In sum, these results indicate that domain-specific numerical precursor skills seem to be more important for children's development of an understanding of the cardinality of small numbers as well as of the concept of zero than the more proximal cross-domain abilities such as language and visuospatial abilities.

Keywords: numerical development, cardinality principle, counting, knower level, visuospatial abilities, language abilities

INTRODUCTION

Counting is an important step in children's numerical development (cf. Fuson, 1988). At an age between 2 and 3 years, children usually start learning the sequence of number words (one, two, three, etc.). In the beginning, children often confuse the sequence of number words but soon learn to recite the number words in the appropriate order. In the present study, we were interested in the

development of early numerical abilities in children with a specific focus on their understanding of cardinality, i.e., their understanding of the fact that each number word represents a specific quantity. In particular, we aimed at identifying relevant predictors of children's early understanding of the cardinality of small numbers. In the context of small numbers, it is interesting to note that children seem to acquire the concept of *zero* in a way that is different from how they acquire the concept of other small numbers. For instance, children hardly ever start counting at *zero*. Thus, the development of children's understanding of the cardinality of small numbers and the concept of *zero* seem to differ. However, while children's understanding of the cardinality of small numbers has been investigated quite well, this is not the case for their acquisition of the concept of *zero*. Therefore, we paid specific attention to the development of children's knowledge of *zero*. In particular, we were interested in whether the same variables that predict children's mastery of the cardinality of small numbers larger than *zero* (i.e., numbers 1–7) also predict the acquisition of the concept of *zero*.

In the following, we first give a brief introduction on the development of children's understanding of the cardinality of small natural numbers and the specificity of the concept of *zero* before going into the details of the current study.

Development of Children's Understanding of Cardinality

Children's first attempts at counting often turn out to be a numerically meaningless recitation of number words (Fuson, 1988). At this stage, children may not yet understand that the number word *two* refers to the numerosity of a set of two objects. In addition to keeping to the correct order of number words, it is also necessary to follow the principle of one-to-one correspondence for successful counting (Gelman and Gallistel, 1978). Finally, to understand that the number word named last actually represents the number of items in the set counted, children need to understand the concept of cardinality.

Most children between 2 and 3 years of age still have trouble in fully understanding cardinality (Fuson, 1988). Interestingly, recent studies have indicated that children do not acquire an understanding of cardinality for all numbers at the same time. Rather, this seems to be a step-by-step process. In the first step, children acquire the cardinal meaning of *one* while all other numbers are simply considered larger than *one* (e.g., Sarnecka and Carey, 2008). When a child at this so-called *one*-knower level (i.e., she/he only understands the cardinal meaning of *one*) is asked to give an experimenter two or three objects, the child will most probably pass more than one object without further differentiating between these larger numerosities. Some months later, children reach the *two*-knower level (i.e., she/he understands the cardinal meaning of *one* and *two*; Sarnecka and Carey, 2008). This level is followed by the *three*-knower level and then by the *four*-knower level, the understanding of each new number being assumed to build on children's understanding of the previous numbers (e.g., Sarnecka et al., 2007). Children at these levels (*one* to *four*) have been termed as "subset-knowers" (Le Corre and Carey, 2007).

After the *five*-knower level has been reached, most children show a change in their further development of understanding the cardinal meaning of number words. Suddenly, they seem to be able to generate the right cardinality for *five* and larger numbers. At this level, children are identified as "cardinality-knowers" (Sarnecka and Carey, 2008). Sarnecka and Carey (2008) explain that cardinality-knowers differ qualitatively from subset-knowers because cardinality-knowers *understand* how counting works. At the age of around three-and-a-half years, children usually master the significance of cardinality by realizing that a set of five objects, labeled with the number word *five*, can also be counted *one, two, three, four, and five* (Mix, 2009).

Importantly, there is accumulating evidence that the above described development of children's understanding of the cardinality of small numbers is influenced by both cross-domain as well as domain-specific numerical abilities (e.g., LeFevre et al., 2010). In particular, the influences of language (e.g., Carey, 2004; Negen and Sarnecka, 2012) as well as visuospatial abilities (e.g., Newcombe et al., 2015 for a review) were observed. Therefore, we specifically considered these two cross-domain abilities when investigating the predictors of children's understanding of the cardinality of small numbers and *zero*. In the following, we will first summarize the evidence of the influence of language and visuospatial abilities on children's understanding of the cardinality of small numbers and *zero* before considering the influences of domain-specific numerical predictors.

Cross-Domain Factors Influencing the Understanding of Cardinality: Language

In this context, the question arises whether there are meaningful predictors of children's understanding of cardinality as sketched above based on knower levels. Recent studies have indicated that language abilities may play an important role. Negen and Sarnecka (2012) found that understanding the cardinal meaning of the first number words was associated with the development of children's vocabulary: the larger a child's vocabulary, the better her/his cardinal number knowledge. Interestingly, this association may be influenced by the fact that linguistic markers might well corroborate differentiating between *one* and *more*. For example, in the English language, any number larger than *one* is usually followed by a plural noun, with "-s" added to the word representing the counted objects (e.g., *one* car but *two* or *more* cars). At such an early stage, these markers may help children differentiate between *one* (singular) and *more* (plural).

This hypothesis is backed by the findings of Sarnecka et al. (2007) who observed that children who speak Japanese, which is a so-called classifier language with no such singular-plural distinction, take longer to understand the cardinality of *one* than English- or Russian-speaking children whose languages differentiate explicitly between singular and plural. Importantly, additional analyses by Sarnecka et al. (2007) indicated that parents from all three countries used number words in a comparable manner when interacting with their children. Moreover, Barner et al. (2007) found that English-speaking children distinguished the quantity of *one* and *more* at about 22–24 months of age. Interestingly, this corresponded to the

same age that parents reported that their children began using plural nouns. Another interesting point is that some languages differ in marking the plural of small magnitudes (2–4) and all magnitudes above *five*. For example, in Slovakian people say: one jablko (apple), two jablka (apples), and five jablk (apples, but with another ending). Thus, it might be assumed that in these languages additional hints are given by these changing plural markers for the important development of children's understanding of the cardinality of small numbers.

This evidence corroborates the claim that language plays a crucial role in children's acquisition of cardinality knowledge of small numbers (Barner et al., 2009). In particular, language was argued to provide essential mental “glue” that enables the human mind to assemble new complex concepts from simple primitives (Spelke, 2003; Condry and Spelke, 2008). This raises the obvious question whether the continuum of numbers is derived from language or whether there are other factors influencing children's early numerical development.

Further Cross-Domain Factors Influencing the Understanding of Cardinality: Visuospatial Abilities

Apart from language abilities, there is also compelling evidence suggesting that visuospatial skills may be associated with children's numerical development (e.g., Ansari et al., 2003; Gunderson et al., 2012; LeFevre et al., 2013; Patro et al., 2014; Pixner et al., 2017 for a review on spatial-numerical association in preliterate children; see Newcombe et al., 2015 for a review on the intertwined development of spatial and numerical skills). This seems reasonable as Newcombe et al. (2015) argue that space and number have a mutual basis, i.e., the generalized magnitude system that is resorted to both simple spatial and numerical tasks. Furthermore, it is supposed that there is a spatial representation of number magnitudes often referred to by the metaphor of a *mental number line*. On the mental number line, numbers are assumed to be represented in ascending magnitude order from left to right (at least in Western countries; e.g., Dehaene et al., 1993). As such, the mental number line represents a combination and integration of spatial and numerical concepts.

Nevertheless, there are only very few studies that pay specific attention to the association between children's visuospatial abilities and their development of cardinal number knowledge during early childhood (e.g., Ansari et al., 2003). Most of the related research examined primary school children and mainly considered the development of the mental number line. Yet, tasks usually employed to assess the mental number line often require both the cardinal knowledge of number magnitudes as well as visuospatial abilities (e.g., the number line estimation task in which a target number has to be located on a number line of which only the start and end points are given, e.g., Siegler and Opfer, 2003).

In this context, Gunderson et al. (2012) observed that in first and second graders, spatial skills predicted the improvement in number line estimation over the course of the school year. In addition, children's spatial skills also predicted later approximate calculation abilities. These findings are substantiated by the

results of training studies in which spatial-numerical trainings were more effective than non-spatial control training in enhancing kindergartners' number line estimation as well as counting performance (Fischer et al., 2011, 2015; Dackermann et al., 2016 for overviews of spatial-numerical trainings). In line with this, Siegler and Ramani (2008) found that knowledge of numerical quantities in 4-year-olds improved significantly when they played board games that involved a physical realization of the mental number line (i.e., moving a token as many steps to the right as there were points on a dice).

Taken together, this suggests that cross-domain abilities such as (visuo-)spatial abilities as well as language abilities (see above) seem to play a crucial role in children's numerical development. Of course, children's numerical development is also influenced considerably by domain-specific numerical predictors as described in the following.

Domain-Specific Basic Numerical Factors Influencing the Understanding of Cardinality

In addition to cross-domain abilities such as language and visuospatial abilities, it was observed that numerical competencies such as children's understanding of cardinality are also influenced by other domain-specific basic numerical competencies such as counting (e.g., Aunola et al., 2004), the ability to identify and name number symbols (e.g., Schmidt, 1982), as well as finger-based numerical representations (e.g., Noel, 2009). While the association is obvious for counting and number identification, the influence of finger-based representations needs a brief introduction. For instance, finger-based number gestures (e.g., thumb, index, and middle finger stretched out to represent three) serve as an important bridge between preverbal mental representation of numbers and number words (e.g., Gunderson et al., 2015; Roesch and Moeller, 2015 for a theoretical discussion). Usually, at the age of two, children begin to use such gestures while counting (Gelman and Gallistel, 1978), exactly at the same time as they begin to understand the cardinality of numbers. This led us to consider finger-based numerical representations when investigating the development of children's understanding of cardinality.

As already mentioned above, a specific focus of the current study was on examining children's understanding of the concept of *zero* by evaluating possible predictors for the acquisition of the concept of *zero*. In particular, we aimed at evaluating whether cross-domain language and/or visuospatial abilities as well as domain-specific numerical factors also play an important role in the acquisition of the concept of *zero*. Or is the mastery of cardinality of small numbers necessary to understand the concept of *zero*?

The Specific Role of Zero

From an evolutionary point of view, *zero* is a rather “young” number (Butterworth, 1999). The use of *zero* was reported first in about 300 BC (Seife, 2000), even though people had used numbers in everyday life long before. To date, only few studies have examined the processing of *zero* and its development in

detail (see Nieder, 2016 for a recent review on the emergence and the development of *zero*). There is evidence that processing *zero* is unique in both children as well as adults (Wellman and Miller, 1986; Brysbaert, 1995). Yet, difficulties in understanding *zero* may not only refer to the numerical value of *zero*, but may originate from difficulties at a more general level of understanding the concept of *nothing*. Wynn and Chiang (1998) analyzed the development of the concept of *no object* in a series of experiments with infants. In these experiments, a single item ‘magically’ appeared/disappeared in a location in which no/an item had been shown before. Infants were not surprised when an object magically appeared. However, they were irritated by the magical disappearance of an object from its former location. From these findings, Wynn and Chiang (1998) concluded that 8-months-old infants were unable to understand *no objects*.

Moreover, Wellman and Miller (1986) reported that children first learn to identify the symbol of *zero* without actually understanding what this symbol means semantically. Only later on, children are assumed to learn that *zero* represents *nothing*, but initially without considering it as a numerical value. Therefore, children at this stage may still not understand whether *zero* is more or less than *one*. At the age of 5 to 6 years, at the end of preschool, however, most children understand that *zero* is a numerical concept and do correctly identify it as the smallest natural number (Wellman and Miller, 1986).

When looking at the development of the differentiation between *one*, *two*, and so on as described above, it becomes clear that *zero* is unique. Interestingly, from a linguistic point of view, *zero* is associated with using the plural form of the respective noun in many languages (e.g., *zero cars* in English, null Autos in German, etc.), even though *zero* is even less than *one* and is found to the left of *one* on the mental number line. Moreover, *zero* is usually not part of children’s common counting sequence. Mostly, children start counting at *one* and not at *zero*. Moreover, unlike other integers, *zero* does not represent the presence of a quantity, but its absence. Accordingly, these specificities may influence children’s understanding of the cardinal meaning of *zero*. For instance, in a magnitude comparison, four-year-old children were just as likely to indicate that *zero* is larger than *three* as vice versa (Merritt and Brannon, 2013). This was examined in a non-symbolic numerosity comparison task, in which trials with *no* objects were presented. Children had to decide, on which one of two pictures they could see more objects. From this result, Merritt and Brannon (2013) concluded that *zero* is represented on the same numerical continuum as other natural numbers at the age of about 4 years.

However, not only for children but also for adults the representation of *zero* seems to be different from the representation of other small numbers. For instance, Brysbaert (1995) found that reading times for small integers (e.g., one, two, or three) were significantly shorter than the reading time for *zero*. This indicates that processing of *zero* differs substantially from processing of other integers and might be based on other principles (Brysbaert, 1995). Grounded on this and other findings, Pinhas and Tzelgov (2012) concluded that *one* may be considered the innately smallest number (Leslie et al., 2008),

whereas *zero* represents a later and the smallest culturally acquired number.

Another role of *zero* is its placeholder function in multidigit numbers. Many studies have documented that children, and adults too, have difficulties in understanding the placeholder function of *zero* (Brown, 1981; Crooks and Flockton, 2002). Further problems representing *zero* were found by Wheeler and Feghali (1983) who observed that adults had more problems completing arithmetic problems when at least one *zero* was involved. Wellman and Miller (1986) inferred that these problems originate from the fact that computations with *zero* usually require the correct application of specific rules ($X \text{ times } 0 \text{ is } 0$, but $X \text{ plus } 0 \text{ is } X$) and thus differ from computations involving other natural numbers.

Considering this representational specificity of *zero*, one cannot be sure that language that was supposed (and observed) to predict children’s acquisition of the cardinality of small numbers also predicts children’s understanding of the concept of *zero*. As mentioned above, nouns linked with *zero* are linguistically marked as plural (e.g., *zero cars*) in many languages. Accordingly, children might misinterpret *zero* as representing a quantity larger than *one*. Therefore, we suggest that children refer to other sources of information to correctly understand the concept of *zero*. In particular, visuospatial abilities associated with processing of spatial attributes of the mental number line (i.e., *zero* being smaller and thus located to the left of *one* on the mental number line) or basic numerical abilities, such as understanding the cardinality of small numbers, may be recruited in this process.

Taken together, the present study set out to evaluate the possibly differential association of cross-domain abilities such as language and visuospatial skills of children with their understanding of the cardinality of small numbers as observed in previous studies while also considering the influences of domain-specific basic numerical abilities (i.e., counting, number identification, and finger-based representations). We hypothesized that the influences of domain-specific basic numerical competencies should outweigh those of cross-domain abilities because they allow for a more specific prediction of later numerical skills. However, going beyond previous studies, we were specifically interested in children’s knowledge of *zero* and whether the acquisition of the concept of *zero* is influenced by language, visuospatial, and basic numerical abilities in a way comparable to the cardinality of small natural numbers. As there is only very little research on the development of children’s knowledge of *zero*, it is hard to derive a specific hypothesis. Nevertheless, similar to the case of children’s understanding of the cardinality of small numbers, we would hypothesize that domain-specific numerical predictors should be more important than cross-domain ones.

MATERIALS AND METHODS

Participants and Sample Description

For our study, children were recruited from local public kindergartens around Innsbruck, Austria. Altogether, 65 children (31 boys and 34 girls) were included in this study. Their

age ranged from 4 to 5 years ($M = 4$ years, 4 months, $SD = 3$ months). Most of the children (81.5%) were right-handed. All children attended the kindergarten regularly for at least 1 year, were monolingual native speakers of German, and showed no intellectual or language impairments. Written informed consent was obtained from parents prior to the study and children were asked for verbal assent prior to assessment. The study was approved by the Research Committee for Scientific and Ethical Questions at UMIT and school authorities of the state of Tyrol, Austria.

Procedure and Tasks

Participating children were tested in German in a single one-on-one session in a quiet room in their kindergarten.

The assessment of children's *numerical and counting skills* comprised four tasks:

- (1) Children were asked up to which number they could count. The number sequence formed by each child was transcribed by the experimenter. The task was stopped when a child was obviously uncomfortable about continuing, began to repeat previously used segments, or was not able to continue her/his sequence any further. The largest number for which the counting sequence was correct was considered as the dependent variable.
- (2) To identify children's knower level for the numbers 1–7, as well as their knowledge of *zero*, we used a variant of the Give-N task. As known from previous research, cardinality generalizes for numbers ≥ 5 (Sarnecka and Carey, 2008). Nevertheless, we tested up to the number 7 to check whether it would be the same or not. Quantities were presented randomly and each quantity was presented only once. This was due to our consideration of using Rasch models to analyze the Give-N task for which the repeated presentation of items is not beneficial (see below for the results of the Rasch analyses; Böhner, 2011). Additionally, this made testing sessions shorter and helped keeping children motivated and attentive. Children were first requested to take the respective number of stones (0–7) out of a box. All children who mastered the cardinality of *one* but failed for the cardinality of *two* and more were grouped into knower level 1. Similarly, all children who mastered the cardinality of *one* and *two* were considered to be in knower level 2 and so forth. Knowledge of *zero* was the criterion to group the children in the zero-knower or no-zero-knower groups for the later analysis. Importantly, correct scoring of children's responses to the *zero* item was not trivial because a correct reaction to this item would be doing simply nothing. Hence, whenever a child did not articulate that she/he did nothing on purpose, experimenters were instructed to ask children whether doing nothing/not responding was their answer to this trial. Thereby, we aimed at substantiating evidence on whether or not children understood the meaning of *zero*.
- (3) Furthermore, to assess children's *number identification* abilities, those had to name a numeral that was presented in Arabic form (i.e., 0–7) on a card. Cards were presented

randomly and each of it one time. Correct answers were awarded one point resulting in a maximum score of eight points. Sum scores served as the dependent variable.

- (4) To identify children's finger knowledge, the children were asked to present a different configuration of fingers. Quantities between *zero* and 10 were asked. Each quantity was presented one time and in random order. Any numerically correct finger configuration was accepted as a correct answer irrespective of whether the produced finger pattern showed a canonical or non-canonical pattern with respect to the standard German finger counting routine. Again, correct answers were awarded 1 point with a maximum of 11 points. Sum scores served as the dependent variable.

The order of these numerical tasks was counterbalanced across participants as far as possible to prevent sequence effects.

Additionally, we used the *visual-perception* subtest of the Beery-Buktenica Developmental Test of Visual-Motor Integration (VMI; Beery, 2004) to assess the visuospatial abilities in children. This task focuses on the visual discrimination component that was important to us and not on fine motor skills, which are often assessed in similar studies. In this paper-and-pencil-based test, children had to complete up to 16 geometric forms/patterns representing items with increasing complexity. For each item, children had to decide which out of four shapes presented in a response box below the actual item fitted the one shape shown as the actual stimulus. Visual discrimination is needed to solve these items. For each correctly solved item, children were awarded one point summing up to a maximum of 16 points in this task with sum scores serving as the dependent variable. We used this task as it focuses on visual discrimination and, thus, seemed more appropriate to us as a measure of visuospatial processing compared to tasks on visuo-motor integration (e.g., Corsi block) often used in other studies (e.g., LeFevre et al., 2010). Please note that comparable tasks focusing on visual discrimination were previously used by, for instance, Zhang et al. (2014) pursuing a similar research question.

Moreover, to assess the general *language* abilities, the standardized active vocabulary test (Aktiver Wortschatztest for 3- to 5-year-old children; AWST-R, Kiese-Himmel, 2005) was administered (following the approach of Negen and Sarnecka, 2012, on measuring language abilities). In this test, children have to name visually presented objects (nouns) and activities (verbs). The test material consists of 75 picture cards (51 nouns and 24 verbs). For each correctly named object or activity of the presented scenarios, children were awarded one point. In this test, a maximum of 75 points could be achieved. Sum scores were used as the dependent variable.

RESULTS

Knower Levels

Results of the Give-N task indicated 1 one-knower, 5 two-knowers, 7 three-knowers, 10 four-knowers, and 42 cardinality-knowers (5, 6, and 7; for more details see Sarnecka and Carey,

2008) in our sample. As to the knowledge of *zero*, we found that 33 children already understood the concept of *zero*, whereas 32 did not yet master this concept.

For the first part of the analyses, children were classified into non-zero-knowers and zero-knowers. For the second part of the analyses, children were classified into groups of subset-knowers and cardinality-knowers. All children on the 1- to 4-knower levels were considered subset-knowers, whereas the others were considered cardinality-knowers (for more details see Sarnecka and Carey, 2008).

Subsequent statistical analyses followed a two-step procedure. In the first step, we evaluated the potential differences between non-zero-knowers and zero-knowers as well as subset-knowers and cardinality-knowers with regard to age, language (vocabulary), and visuospatial abilities as well as number identification, finger knowledge, and counting abilities, and the actual knower level if they understood *zero* and accordingly had the knowledge or non-knowledge of *zero* at the actual knower level, using *t*-tests.

In the second step, we conducted regression analyses to evaluate the predictive value of the above mentioned predictors for knowledge of *zero* as well as children's cardinality knowledge, that is whether and which of these competencies are relevant for children's acquisition of the concept of *zero* and the cardinality of small numbers. As regards knowledge of *zero*, we ran a logistic regression analysis predicting zero-knowers vs. non-zero-knowers, whereas a multiple linear regression analysis was conducted to predict children's knower level reflecting their understanding of the cardinality of small numbers (continuously coded for cardinalities from 1 to 4 plus cardinality-knowers).

In both regression analyses, predictors were considered block-wise. In the first block, non-numeric predictors, vocabulary, and visuospatial perception were incorporated in the regression model. In the second block, basic numerical abilities, number identification, finger knowledge, and counting abilities were included in the model. In the last step, the knower level (continuously coded for cardinalities from 1 to 4 plus cardinality-knowers) or knowledge of *zero* (coded categorically 1 or -1 for successful or not successful understanding of *zero*, respectively) was included. A $p < 0.05$ level of significance for the change in R^2 was applied for the inclusion of the predictors in the regression model.

Mastery of the Concept of Zero

The first part of the analyses addressed children's knowledge of *zero* in the present sample. Interestingly, 14 out of the 42 cardinality-knowers did not show understanding of *zero*, whereas there were 5 out of 23 subset-knowers who already understood the concept of *zero*. This descriptive analysis shows that there might be a double dissociation between understanding the cardinality of small numbers and understanding the concept of *zero* as there are children in our sample who have already acquired one concept but not the other one or vice versa.

These first indications for differences in children's understanding of the concept of *zero* and the cardinality of small numbers were substantiated by an analysis of the discrimination of respective items; that means, the item measuring the concept

of *zero* may allow for differing discrimination compared to the items for small numbers. The hypothesis of equal item discrimination can be tested in the Rasch model (Rasch, 1960) by applying the so-called pseudo-exact or conditional tests (Ponocny, 2001; Draxler and Zessin, 2015), which are particularly suited for small sample sizes. The results of the conditional tests yielded a p -value of 0.044 for the item measuring the concept of *zero* and considerably higher p -values for the rest of the items, indicating that *zero* seems to be processed differently. These results are in accordance with the descriptive analysis (see above). Furthermore, a Bayesian analysis according to Draxler (2018) substantiated these results. The obtained posterior distributions indicated the item assessing the understanding of the concept of *zero* as the one with deviating discrimination in comparison to the other items.

Second, we evaluated the differences between non-zero-knowers and zero-knowers in terms of age, vocabulary, visuospatial perception, number identification, counting abilities, finger knowledge, and knower level (cf. Table 1). We observed that zero-knowers were significantly better than non-knowers of *zero* at vocabulary, counting abilities, number identification, finger knowledge, and knower level, but not on visuospatial perception.

Because no age differences between groups were observed and no significant correlation between age and knowledge of *zero* ($r = -0.05$, $p = 0.668$, see Table 2 for correlations of other variables), age was no longer considered in the regression.

In the next step, a logistic regression analysis with knowledge regarding *zero* (successful vs. not successful) as the dependent variable was run. In the first block, we included the predictors visuospatial perception and vocabulary. The model showed a significant goodness of fit [$\chi^2(2) = 11.55$, $p = 0.003$] with a Cox and Snell R^2 value of 0.16 and a Nagelkerke's pseudo R^2 value of 0.22, which corresponds to a strong effect according to Cohen (1992). Only vocabulary turned out to be a significant predictor with better vocabulary predicting better *zero* knowledge ($b = 0.098$, $SE = 0.32$, odds ratio = 1.103, $p = 0.003$).

In the second block, counting abilities, number identification, and finger knowledge were included in the analysis as additional predictors. Again, the model fit the data significantly [$\chi^2(5) = 24.58$, $p < 0.001$, Cox and Snell $R^2 = 0.32$, pseudo $R^2 = 0.43$, which corresponds to a very strong effect according to Cohen (1992)]. Here, only counting abilities ($b = 0.199$, $SE = 0.084$, odds ratio = 1.22, $p = 0.017$) were a significant predictor. Inspection of beta weights indicated that better *zero* knowledge was associated with better counting abilities. Finger knowledge, number identification, visuospatial perception, and vocabulary did not account for a significant part of the variance.

In the third block, children's knower level was included in the model. The final model again fit the data significantly well [$\chi^2(6) = 26.44$, $p < 0.001$, Cox and Snell $R^2 = 0.34$, pseudo $R^2 = 0.46$, corresponding to a very strong effect according to Cohen (1992)]. Again, only counting abilities ($b = 0.178$, $SE = 0.085$, odds ratio = 1.19, $p = 0.036$) were found to be a significant predictor of *zero* knowledge. Better counting abilities were associated with better *zero* knowledge. Finger knowledge, number identification, visuospatial perception, vocabulary, and

knower level were not considered as meaningful for identifying *zero*-knowers in our sample.

Children's Understanding of Cardinality

As can be read from **Table 3**, regarding knower levels we found significant differences between cardinality-knowers and subset-knowers for the vocabulary task, the number identification task, the counting ability task, the finger knowledge task, and knowledge of *zero*. Cardinality-knowers showed higher scores as compared to subset-knowers on all of these tasks. No significant differences were found for visuospatial abilities and with regard to age.

Because no age differences between the groups were found and the correlation between age and knower level was not significant ($r = -0.135$, $p = 0.282$, see **Table 2** for the correlation matrix of predictors), age was no longer considered in the regression analyses.

In the next step, a multiple linear regression analysis was conducted to predict children's knower level reflecting their understanding of the cardinality of small numbers (continuously coded for cardinalities from 1 to 4 plus cardinality-knowers). In the first block, we included the predictors visuospatial abilities and vocabulary. Only vocabulary accounted for a significant part of the variance [$R^2 = 0.17$, $adj. R^2 = 0.15$, $F(1, 63) = 12.57$, $p = 0.001$]. Inspection of beta weights indicated that increases in vocabulary (constant = 2.465; $B = 0.08$, $SE = 0.02$, standardized $\beta = 0.408$; $p = 0.001$) were associated with a higher knower level.

In the second block, additionally, counting abilities, number identification, and finger knowledge were considered as predictors in the analysis. In the final model [$R^2 = 0.47$, $adj.$

$R^2 = 0.46$, $F(2, 58) = 26.16$, $p < 0.001$], the predictors counting abilities (constant = 1.485; $B = 0.10$, $SE = 0.03$, standardized $\beta = 0.316$; $p = 0.003$) and finger knowledge ($B = 0.60$, $SE = 0.12$, standardized $\beta = 0.506$; $p < 0.001$) accounted for a significant part of the variance. Inspection of beta weights indicated that better counting abilities and higher finger knowledge were associated with higher knower level. In contrast, number identification, visuospatial perception, and vocabulary did not account for a significant part of the variance. Please also note that vocabulary was no longer a significant predictor of knower level as soon as either counting abilities or finger knowledge was considered in the model.

In the third block, knowledge of *zero* was included in the model as an additional predictor. However, this did not change the predictors considered in the final regression model. This indicated that children's knowledge of *zero* did not seem to be predictive of their cardinal number knowledge of small numbers.

DISCUSSION

The present study aimed at investigating possibly differential prediction of cross-domain abilities such as language skills and visuospatial abilities as well as domain-specific abilities such as counting, finger knowledge, and number identification skills of kindergartners' understanding of the concept of *zero* and the cardinality of small numbers. In the following, we will elaborate on these points in turn.

As we were particularly interested in the development of the concept of *zero*, our first objective was to identify predictors for

TABLE 1 | Statistical details of the comparison of non-knowers and knowers of *zero*.

	Non-zero-knowers $N = 32$	Zero-knowers $N = 33$		
	Mean (SE)	Mean (SE)	t	p
Age	52.47 (0.58)	52.15 (0.45)	$t(63) = 0.43$	0.668
Knower level of numbers	4.44 (0.34)	6.30 (0.23)	$t(63) = 4.56$	<0.001
Visuospatial perception (VMI)	10.50 (0.47)	10.21 (0.45)	$t(63) = 0.44$	0.661
Vocabulary (AWST)	33.75 (1.55)	41.36 (1.64)	$t(63) = 3.37$	0.001
Counting abilities	9.87 (0.83)	15.88 (0.96)	$t(61^*) = 4.69$	<0.001
Number identification	3.22 (0.49)	5.30 (0.40)	$t(63) = 3.29$	0.002
Finger knowledge	3.81 (0.27)	4.97 (0.24)	$t(63) = 3.20$	0.002

SE of the mean given in parenthesis. *Two children refused this task, but results did not change marginally when children were omitted from all analyses.

TABLE 2 | Correlation of all variables.

Task	1	2	3	4	5	6	7
1. Vocabulary	1						
2. Visuospatial perception	0.13	1					
3. Number identification	0.37**	0.11	1				
4. Finger knowledge	0.26*	0.13	0.63**	1			
5. Counting ability	0.33**	-0.27*	0.45**	0.39**	1		
6. Knower level	0.41**	0.03	0.57**	0.65**	0.52**	1	
7. Knowledge of <i>zero</i>	0.39**	-0.06	0.39**	0.36**	0.51**	0.50**	1

* $p < 0.05$; ** $p < 0.01$.

TABLE 3 | Statistical details of the comparisons between subset-knowers and cardinality-knowers.

	Subset-knowers <i>N</i> = 23	C-knowers <i>N</i> = 42		
	Mean (SE)	Mean (SE)	<i>t</i>	<i>p</i>
Age	53.09 (0.67)	51.88 (0.42)	<i>t</i> (63) = 1.60	0.115
Visuospatial perception (VMI)	10.43 (0.55)	10.31 (0.41)	<i>t</i> (63) = 0.18	0.855
Vocabulary (AWST)	32.96 (1.70)	40.17 (1.51)	<i>t</i> (63) = 3.00	0.004
Counting abilities	9.19 (1.15)	14.93 (0.81)	<i>t</i> (61*) = 4.09	<0.001
Number identification	2.52 (0.52)	5.24 (0.37)	<i>t</i> (63) = 4.31	<0.001
Finger knowledge	3.17 (0.29)	5.07 (0.18)	<i>t</i> (63) = 5.75	<0.001

SE of the Mean given in parenthesis. *Two children refused this task, results did not change when these children were omitted from all analyses.

children's understanding of the concept of *zero*. The results of the regression analyses as well as the Rasch analysis showed a significant difference between the understanding of cardinality of small numbers and the concept of *zero*. Descriptive analyses also showed that 14 out of the 42 cardinal-principle-knowers did not show understanding of *zero* whereas 5 out of 23 subset-knowers already understood the concept of *zero*. This provides further evidence for the claim that cardinality-knowledge for small numbers and *zero* seems to develop differently. Therefore, we assumed that different processes might be responsible for the development of these two concepts.

Therefore, we first evaluated whether there were group differences on cross-domain (i.e., language and visuospatial abilities) as well as domain-specific numerical variables (i.e., counting skills, number identification, finger knowledge, and knower level) between children who already mastered the concept of *zero* and those (children) who did not. Results indicated the expected significant differences between the two groups in language, counting abilities, number identification, finger knowledge, and children's knower level. Children who already mastered the concept of *zero* showed better performance on all of the respective abilities, but not with regard to visuospatial abilities.

To evaluate the predictive value of cross-domain (i.e., language and visuospatial abilities) and domain-specific (i.e., counting, number identification, and finger knowledge) variables for children's understanding of the concept of *zero*, we followed a three-stage procedure with logistic regression analyses. We first incorporated cross-domain variables and observed that language, but not visuospatial abilities, was a relevant predictor for children's understanding of the concept of *zero*. When considering counting, number identification, and finger knowledge in the second step, only counting skills remained as a significant predictor of children's understanding of *zero*. Finally, in the third step, the significant influence of counting abilities was prevailing when considering children's knower level.

These results were only partially in line with our expectations. On the one hand, we found that in the final model, language did not account for a unique part of variance in children's understanding of the concept of *zero*. In German (i.e., the first language of the children examined in the current study) as well as in English and many other languages, *zero* as a number is followed by the plural form of a noun (e.g., *zero cars*). While the plural form correctly indicates the differentiation between

one and *more* and may thus help children acquire the cardinality principle of small numbers, it may be a hurdle for children's acquisition of the concept of *zero*. In line with this notion, we did observe that language was no longer a significant predictor of children's understanding of *zero* as soon as domain-specific numerical variables were considered in the model. In particular, only counting skills were found to be of a significant predictive value for children's understanding of the concept of *zero*. However, language, but not visuospatial abilities, was a significant predictor when only cross-domain variables were considered. This finding is hard to reconcile with the notion that because of the inconsistencies regarding its language coding, *zero* might rather be internalized by visuospatial representations. In sum, our findings suggest that language in general and language-based specific numerical skills such as counting seem to be significant predictors of children's early understanding of the concept of *zero*. Thus, these findings indicate that the understanding of the cardinality of small numbers and the concept of *zero* seem to be rather independent of each other.

Apart from investigating the predictors of children's early understanding of the concept of *zero*, we were interested in children's understanding of the cardinality in the number range from *one* to *seven*. This was motivated by the findings of Sarnecka and Carey (2008) who claimed significant differences between subset-knowers and cardinality-knowers; that means, children who only internalized the cardinality for a subset of numbers (e.g., 1-, 2-, or 3-knowers) and children who already understood the cardinality of numbers up to *five* and beyond. Our analyses substantiated the expected differences between the two groups in language (vocabulary), counting abilities, number identification, and finger knowledge, but again not with regard to visuospatial abilities.

Comparable to the case of children's understanding of the concept of *zero*, we then ran regression analyses to evaluate the predictive value of cross-domain (i.e., language and visuospatial abilities) and domain-specific (i.e., counting, number identification, finger knowledge) variables as well as the influence of children's understanding of the concept of *zero*. In the first step, we considered only cross-domain variables in the regression analysis. We observed that language, but not visuospatial skills, was a relevant predictor for children's cardinality knowledge. In the next step, we further included domain-specific numerical predictors and found that this led to the observation that language was no longer a significant predictor of children's

understanding of the cardinality of small numbers. Instead, the latter was predicted significantly by children's counting abilities as well as their finger number knowledge. This is in line with the findings of Reeve and Humberstone (2011) who found a positive association between children's use of finger-based numerical representations and their early arithmetic competencies. Additional consideration of children's knowledge of *zero* in the third step did not improve the regression model. Thus, results indicated that children's understanding of the cardinality of small numbers might be associated with their language (vocabulary) skills. However, as soon as more domain-specific predictors were considered, the latter (i.e., counting skills and finger knowledge) seemed to overrule the influence of language.

As regards the relevant predictors, these findings are in line with earlier findings (e.g., Fuson, 1988; Gunderson et al., 2015) arguing for the importance of domain-specific abilities for the acquisition of the cardinality of small numbers. Moreover, our findings at least partly fit those of Negen and Sarnecka (2012), suggesting a more general influence of language skills on children's understanding of cardinality (as reflected by the significant zero-order correlation of vocabulary and both children's *zero* knowledge as well their understanding of cardinality). However, in the present study, visuospatial abilities were not a relevant predictor for children's understanding of the cardinality of small numbers. This seems to be in contrast to the findings of Pixner et al. (2017) who observed an association between visuospatial abilities and basic numerical competencies in kindergartners. Yet, a closer look at the study reveals at least two possible reasons for these differences. First, in the present study we specifically focused on investigating children's understanding of the cardinality of small numbers, whereas Pixner et al. (2017) measured a broader concept of basic numerical abilities. Second, these differential results may be related to the age of the present sample. Children in our study were on average 1.6 years younger than those assessed in the study of Pixner et al. (2017). Therefore, one might speculate that visuospatial abilities only gain significance for numerical development at a later point in time. There is tentative evidence corroborating this hypothesis. On the one hand, the children examined in studies suggesting the influences of visuospatial abilities on numerical development (as described in more detail in the introduction, e.g., Siegler and Booth, 2004; Gunderson et al., 2012) were again older than the children of the present sample. Additionally, a recent review indicated that the associations of visuospatial and numerical representations become more pronounced with increasing age (McCrink and Opfer, 2014; Newcombe et al., 2015, for a review on the intertwined development of spatial and numerical competencies). In sum, this asks for future longitudinal studies evaluating specifically the interrelations and differential influences between visuospatial abilities and basic numerical competencies in children's cognitive development.

Importantly, the present study only represents a first step toward a better understanding of children's mastery of the concept of *zero*. Future studies are needed to further increase our knowledge on the acquisition of this important concept. An

avenue for such studies may be to consider indefinite numeric quantifiers such as *none* and *nothing* and to evaluate their role in the acquisition of the concept of *zero*. Children in kindergarten may more likely be faced with the words *none* and *nothing* than with *zero* and need to integrate and combine these constructs with their concept of *zero*. In this context, it would be desirable to conduct multiple assessments of the understanding of *zero* but also of the cardinality of the numbers 1–7 to increase the reliability of the measures. Additionally, visuospatial abilities may not be considered a unitary construct but seen to involve several subskills and processes (e.g., Ansari et al., 2003). As such, it is certainly premature to suggest that visuospatial abilities are unrelated to numerical abilities (cf. Newcombe et al., 2015). Instead, it would be interesting to assess the different aspects of visuospatial abilities in future studies to better understand which aspects are and which are not related to numerical abilities, in more detail.

Finally, it is important to not overstate the observed non-significant influences of vocabulary and visuospatial abilities in our final regression models as suggesting that these variables would not be important for children's numerical development in general and their understanding of the concept of *zero* as well as the cardinality of small numbers in particular. There is considerable evidence for the critical influence of these variables (e.g., Ansari et al., 2003; Barner et al., 2009; Negen and Sarnecka, 2012) and we observed a significant influence of vocabulary on both children's understanding of the concept of *zero* as well as the cardinality of small numbers before considering domain-specific numerical predictors in the regression models. As such, it seems a question of proximity between the predictor and the criterion variable that needs to be considered.

In other words, when controlling a predictor (e.g., vocabulary) for a more proximal variable (e.g., counting abilities) makes the predictor non-significant, that does not necessarily mean it is not an important causal predictor of the criterion variable (in this case *zero* knowledge and understanding of the cardinality of small numbers). Even though vocabulary and visuospatial abilities may not be considered as immediate proximal causes of the learning of number words or the acquisition of the concept of *zero*, the influences of these cross-domain variables exist at different levels of how we conceptualize the learning process. Counting ability, for instance, may be considered a direct prerequisite for acquiring cardinality knowledge. This sets it as a more proximal and direct cause, which needs to be dealt with in a somewhat separate way from the broader aspects of cognitive development like general vocabulary and visuospatial abilities.

Therefore, our argumentation is not about downplaying the influences of less proximal cross-domain cognitive abilities on children's numerical development. However, evaluating the influences of broader cross-domain and proximal domain-specific variables as well as their potential interplay would require a longitudinal dataset for which direct versus indirect effects of the respective predictors as well as potential mediating effects can be evaluated. Therefore, future longitudinal studies would be desirable that not only consider more specific aspects of cross-domain abilities but also allow the evaluation of the direct as well as indirect influences of

proximal domain-specific numerical and broader cross-domain variables as well as their interplay.

CONCLUSION AND PERSPECTIVES

The present study aimed at evaluating the differential influences of cross-domain abilities (i.e., language and visuospatial skills) and domain-specific basic numerical abilities (i.e., counting, number identification, and finger-based representations) on kindergartners' understanding of the concept of zero and the cardinality of small numbers. In sum, our results indicated that children's understanding of both the concept of zero and the cardinality of small numbers was associated significantly with their language skills. However, this association became insignificant as soon as domain-specific numerical predictors were considered. This substantiates the relevance of basic numerical competencies for children's early numerical development. However, as discussed above, the present study could not identify whether the relevance of cross-domain and domain-specific variables for children's numerical development differs over time. It might be that in some periods (for instance, during early numerical development in kindergarten), domain-specific numerical competencies are specifically important when children need to build up an abstract knowledge of number magnitudes. As indicated by the use of fingers for counting and initial arithmetic, building up this knowledge may be bound

more closely to domain-specific aspects. Later on, when children will have successfully understood the cardinality of number magnitudes, cross-domain abilities may gain influence (e.g., Geary et al., 2017). As such, future longitudinal studies on children's early numerical development in kindergartens would be desirable to evaluate these claims.

ETHICS STATEMENT

RCSEQ, Research Committee for Scientific and Ethical Questions from the UMIT, Hall in Tyrol. This study was carried out in accordance with the recommendations of 'RCSEQ, Research Committee for Scientific and Ethical Questions from the UMIT, Hall in Tyrol' with written informed consent from all subjects. All subjects gave written informed consent in accordance with the Declaration of Helsinki. The protocol was approved by the 'RCSEQ, Research Committee for Scientific and Ethical Questions from the UMIT, Hall in Tyrol.'

AUTHOR CONTRIBUTIONS

SP substantially contributed to the conception or design of the work and the acquisition, analysis, or interpretation of data for the work. VD drafted a part of the work. KM drafted a part of the work and revised it critically for important intellectual content.

REFERENCES

- Ansari, D., Donlan, C., Thomas, M. S. C., Ewing, S. A., Peen, T., and Karmiloff-Smith, A. (2003). What makes counting count? Verbal and visuo-spatial contributions to typical and atypical number development *J. Exp. Child Psychol.* 85, 50–62. doi: 10.1016/S0022-0965(03)00026-2
- Aunola, K., Leskinen, E., Lerkkanen, M. K., and Nurmi, J. E. (2004). Developmental dynamics of math performance from preschool age to grade 2. *J. Educ. Psychol.* 96, 699–713. doi: 10.1037/0022-0663.96.4.699
- Barner, D., Chow, K., and Yang, S.-J. (2009). Finding one's meaning: a test of the relation between quantifiers and integers in language development. *Cogn. Psychol.* 58, 195–219. doi: 10.1016/j.cogpsych.2008.07.001
- Barner, D., Thalwitz, D., Wood, J., Yang, S.-J., and Carey, S. (2007). On the relation between the acquisition of singular-plural morpho-syntax and the conceptual distinction between one and more than one. *Dev. Sci.* 10, 365–373. doi: 10.1111/j.1467-7687.2007.00591.x
- Beery, K. E. (2004). *The Beery-Buktenica Developmental Test of Visual-Motor Integration*, 5th Edn, Minneapolis, MN: NCS Pearson, Inc.
- Brown, M. (1981). "Number Operations," in *Children's Understanding of Mathematics*, ed. K. Hart (London: John Murray), 23–47.
- Brysbaert, M. (1995). Arabic number reading: on the nature of the numerical scale and the origin of phonological recording. *J. Exp. Psychol. Gen.* 122, 434–452. doi: 10.1037/0096-3445.124.4.434
- Bühner, M. (2011). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson Studium.
- Butterworth, B. (1999). *The Mathematical Brain*. London: Macmillan.
- Carey, S. E. (2004). Bootstrapping & the origin of concepts. *Daedalus* 133, 59–68. doi: 10.1162/001152604772746701
- Cohen, J. (1992). Statistical power analysis. *Curr. Dir. Psychol. Sci.* 1, 98–101. doi: 10.1111/1467-8721.ep10768783
- Condry, K. F., and Spelke, E. S. (2008). The development of language and abstract concepts: the case of natural number. *J. Exp. Psychol. Gen.* 137, 22–38. doi: 10.1037/0096-3445.137.1.22
- Crooks, T., and Flockton, L. (2002). *Mathematics Assessment Results 2001. National Education Monitoring*. Dunedin: Educational Assessment Research Unit.
- Dackermann, T., Fischer, U., Cress, U., Nuerk, H.-C., and Moeller, K. (2016). Bewegtes lernen numerischer kompetenzen. Themenheft: psychologie und Wissensmedien. *Psychol. Rundsch.* 67, 102–109. doi: 10.1026/0033-3042/a000302
- Dehaene, S., Bossini, S., and Giraux, P. (1993). The mental representation of parity and number magnitude. *J. Exp. Psychol.* 122, 371–396. doi: 10.1037/0096-3445.122.3.371
- Draxler, C. (2018). Bayesian conditional inference for Rasch models. *Adv. Stat. Anal.* 102, 245–262. doi: 10.1007/s10182-017-0303-6
- Draxler, C., and Zessin, J. (2015). The power function of conditional tests of the Rasch model. *Adv. Stat. Anal.* 99, 367–378. doi: 10.1007/s10182-015-0249-5
- Fischer, U., Link, T., Cress, U., Nuerk, H.-C., and Moeller, K. (2015). "Math with the dance mat: on the benefits of embodied numerical training approaches," in *Learning Technologies and the Body: Integration and Implementation in Formal and Informal Learning Environments*, ed. V. Lee (New York, NY: Routledge), 149–163.
- Fischer, U., Moeller, K., Bientzle, M., Cress, U., and Nuerk, H.-C. (2011). Sensori-motor spatial training of number magnitude representation. *Psychol. Bull. Rev.* 18, 177–183. doi: 10.3758/s13423-010-0031-3
- Fuson, K. C. (1988). *Children's Counting and Concepts of Number*. New York, NY: Springer Verlag. doi: 10.1007/978-1-4612-3754-9
- Geary, D. C., Nicholas, A., Li, Y., and Sun, J. (2017). Developmental change in the influence of domain-general abilities and domain-specific knowledge on mathematics achievement: an eight-year longitudinal study. *J. Educ. Psychol.* 109, 680–693. doi: 10.1037/edu0000159
- Gelman, R., and Gallistel, C. R. (1978). *The Child's Understanding of Number*. Cambridge, MA: Harvard University Press.
- Gunderson, E. A., Ramirez, G., Beilock, S. L., and Levine, S. C. (2012). The relation between spatial skill and early number knowledge: the role of the linear number line. *Dev. Psychol.* 48, 1229–1241. doi: 10.1037/a0027433

- Gunderson, E. A., Spaepen, E., Gibson, D., Goldin-Meadow, S., and Levine, S. C. (2015). Gesture as a window onto children's number knowledge. *Cognition* 144, 14–28. doi: 10.1016/j.cognition.2015.07.008
- Kiese-Himmel, C. (2005). *AWST-R. Aktiver Wortschatztest für 3- bis 5-Jährige Kinder – Revision*. Göttingen: Hogrefe Verlag.
- Le Corre, M., and Carey, S. (2007). One, two, three, four, nothing more: an investigation of the conceptual sources of the verbal counting principles. *Cognition* 105, 395–438. doi: 10.1016/j.cognition.2006.10.005
- LeFevre, J.-A., Fast, L., Skwarchuk, S.-L., Smith-Chant, B. L., Bisanz, J., Kamawar, D., et al. (2010). Pathways to mathematics: longitudinal predictors of performance. *Child Dev.* 81, 1753–1767. doi: 10.1111/j.1467-8624.2010.01508.x
- LeFevre, J.-A., Jiménez Lira, C., Sowinski, C., Cankaya, O., Kamawar, D., and Skwarchuk, S.-L. (2013). Charting the role of the number line in mathematical development. *Front. Psychol.* 4:641. doi: 10.3389/fpsyg.2013.00641
- Leslie, A. M., Gelman, R., and Gallistel, C. R. (2008). The generative basis of natural number concepts. *Trends Cogn. Sci.* 12, 213–218. doi: 10.1016/j.tics.2008.03.004
- McCrink, K., and Opfer, J. E. (2014). Development of spatial-numerical associations. *Curr. Dir. Psychol. Sci.* 23, 439–445. doi: 10.1177/0963721414549751
- Merritt, D. J., and Brannon, E. M. (2013). Nothing to it: precursors to a zero concept in preschoolers. *Behav. Process.* 93, 91–97. doi: 10.1016/j.beproc.2012.11.001
- Mix, K. S. (2009). How spencer made number. First uses of the number words. *J. Exp. Child Psychol.* 102, 427–444. doi: 10.1016/j.jecp.2008.11.003
- Negen, J., and Sarnecka, B. W. (2012). Number-concept acquisition and general vocabulary development. *Child Dev.* 83, 2019–2027. doi: 10.1111/j.1467-8624.2012.01815.x
- Newcombe, N. S., Levine, S. C., and Mix, K. S. (2015). Thinking about quantity: the intertwined development of spatial and numerical cognition. *Wiley Interdiscip. Rev. Cogn. Sci.* 6, 491–505. doi: 10.1002/wcs.1369
- Nieder, A. (2016). Review: representing something out of nothing: the dawning of zero. *Trends Cogn. Sci.* 20, 830–842. doi: 10.1016/j.tics.2016.08.008
- Noel, M.-P. (2009). Finger gnosis: a predictor of numerical abilities in children? *Child Neuropsychol.* 11, 413–430.
- Patro, K., Nuerk, H.-C., Cress, U., and Haman, M. (2014). How number-space relationships are assessed before formal schooling: a taxonomy proposal. *Front. Psychol.* 5:419. doi: 10.3389/fpsyg.2014.00419
- Pinhas, M., and Tzelgov, J. (2012). Expanding on the mental number line: zero is perceived as the “Smallest”. *J. Exp. Psychol. Learn. Mem. Cogn.* 38, 1187–1205. doi: 10.1037/a0027390
- Pixner, S., Kraut, C., and Dresen, V. (2017). Early predictors for basic numerical and magnitude competencies in preschool children – are they the same or different regarding specific subgroups? *Psychology* 8, 271–286. doi: 10.4236/psych.2017.82016
- Ponocny, I. (2001). Nonparametric goodness-of-fit tests for the rasch model. *Psychometrika* 66, 437–459. doi: 10.1007/BF02294444
- Rasch, G. (1960). *Probabilistic Models for Some Intelligence and Attainment Tests*. Copenhagen: Danish Institute for Educational Research.
- Reeve, R., and Humberstone, J. (2011). Five- to 7-years-olds' finger gnosis and calculation abilities. *Front. Psychol.* 2:359. doi: 10.3389/fpsyg.2011.00359
- Roesch, S., and Moeller, K. (2015). Considering digits in a current model of numerical development. *Front. Hum. Neurosci.* 8:1062. doi: 10.3389/fnhum.2014.01062
- Sarnecka, B. W., and Carey, S. (2008). How counting represents number: what children must learn and when they learn it. *Cognition* 108, 662–674. doi: 10.1016/j.cognition.2008.05.007
- Sarnecka, B. W., Kamenskaya, V. G., Yamana, Y., Ogura, T., and Yudovina, J. B. (2007). From grammatical number to exact numbers: early meanings of “one,” “two,” and “three” in English. *Cogn. Psychol.* 55, 136–168. doi: 10.1016/j.cogpsych.2006.09.001
- Schmidt, R. (1982). Ziffernkenntnis und Ziffernverständnis der Schulanfänger. *Grundschule* 14, 166–167.
- Seife, C. (2000). *Zero: The Biography of a Dangerous Idea*. New York, NY: Penguin Books.
- Siegler, R. S., and Booth, J. L. (2004). Development of numerical estimation in young children. *Child Dev.* 15, 428–444. doi: 10.1111/j.1467-8624.2004.00684.x
- Siegler, R. S., and Opfer, J. E. (2003). The development of numerical estimation: evidence for multiple representations of numerical quantity. *Psychol. Sci.* 14, 237–243. doi: 10.1111/1467-9280.02438
- Siegler, R. S., and Ramani, G. B. (2008). Playing linear numerical board games promotes low-income children's numerical development. *Dev. Sci. Special Issue Math. Cogn.* 11, 655–661. doi: 10.1111/j.1467-7687.2008.00714.x
- Spelke, E. S. (2003). “What makes us smart? Core knowledge and natural language,” in *Language in Mind*, eds D. Genter and S. Goldin-Meadow (Cambridge, MA: MIT Press), 277–311.
- Wellman, H. M., and Miller, K. F. (1986). Thinking about nothing: development of concepts of zero. *Br. J. Dev. Psychol.* 4, 31–42. doi: 10.1016/j.tics.2016.08.008
- Wheeler, M. M., and Feghali, I. (1983). Much ado about nothing: preservice elementary school teachers' concept of zero. *J. Res. Math. Educ.* 14, 147–155. doi: 10.2307/748378
- Wynn, K., and Chiang, W.-C. (1998). Limits to Infants'. *Psychol. Sci.* 9, 448–455. doi: 10.1111/1467-9280.00084
- Zhang, X., Koponen, T., Räsänen, P., Aunola, K., Lerkkanen, M.-K., and Nurmi, J.-E. (2014). Linguistic and spatial skills predict early arithmetic development via counting sequence knowledge. *Child Dev.* 85, 1091–1107. doi: 10.1111/cdev.12173

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Pixner, Dresen and Moeller. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Variability in the Alignment of Number and Space Across Languages and Tasks

Andrea Bender^{1,2*}, Annelie Rothe-Wulf³ and Sieghard Beller^{1,2}

¹ Department of Psychosocial Science, University of Bergen, Bergen, Norway, ² SFF Centre for Early Sapiens Behaviour (SapienCE), University of Bergen, Bergen, Norway, ³ Department of Psychology, Freiburg University, Freiburg, Germany

OPEN ACCESS

Edited by:

Krzysztof Cipora,
Universität Tübingen, Germany

Reviewed by:

Silvia Pixner,
UMIT - Private Universität für
Gesundheitswissenschaften,
Medizinische Informatik und Technik,
Austria
Christine Schiltz,
University of Luxembourg,
Luxembourg

*Correspondence:

Andrea Bender
andrea.bender@uib.no

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 31 January 2018

Accepted: 27 August 2018

Published: 04 October 2018

Citation:

Bender A, Rothe-Wulf A and Beller S
(2018) Variability in the Alignment of
Number and Space Across
Languages and Tasks.
Front. Psychol. 9:1724.
doi: 10.3389/fpsyg.2018.01724

While the domains of space and number appear to be linked in human brains and minds, their conceptualization still differs across languages and cultures. For instance, frames of reference for spatial descriptions vary according to task, context, and cultural background, and the features of the mental number line depend on formal education and writing direction. To shed more light on the influence of culture/language and task on such conceptualizations, we conducted a large-scale survey with speakers of five languages that differ in writing systems, preferences for spatial and temporal representations, and/or composition of number words. Here, we report data obtained from tasks on ordered arrangements, including numbers, letters, and written text. Comparing these data across tasks, domains, and languages indicates that, even within a single domain, representations may differ depending on task characteristics, and that the degree of cross-domain alignment varies with domains and culture.

Keywords: number, space, space-number mapping, mental number line, frames of reference, culture, language

INTRODUCTION

It has long been proposed that humans tend to represent abstract domains such as time or number in terms of more concrete domains such as space (Lakoff and Johnson, 1980). Indeed, evidence for this cross-domain mapping has accumulated over the past 25 years (e.g., Boroditsky, 2000; Fischer and Fias, 2005; Núñez and Cooperrider, 2013). Temporal sequences and events, for instance, appear to be represented along a spatially extending *mental time line* (MTL), as attested to both in linguistic and non-linguistic tasks (overview in Bonato et al., 2012; Bender and Beller, 2014). Likewise, numbers appear to be represented along a spatially extending *mental number line* (MNL), as attested to in tasks using both explicit and implicit measures, such as those concerned with number line estimations (Siegler and Opfer, 2003; Moeller et al., 2009) or with the *spatial-numerical association of response codes* (SNARC) effect (Dehaene et al., 1993; Wood et al., 2008). MTL and MNL have in common that they are assumed to extend in a more or less spatial manner, along one dimension, in one direction, and potentially ad infinitum. An increasing body of evidence related to these constructs seems to corroborate that the domains of space, time, and number are intrinsically linked in human minds, and perhaps even in human brains (Walsh, 2003).

Yet, some observations appear to be at odds with such linear representations, pointing to the possibility that these representations might be neither innate nor universal (e.g., Núñez, 2008, 2011; Bender and Beller, 2014). In particular, three sets of findings are inconsistent with a simply painted picture of cross-domain congruency: (1) the remarkable degree of variability in representations, both within and across domains; (2) the deep impact of cultural practices on the shape of these

representations; and (3) the dependency of such representations on task specifics and context.

Variability in Representations

Time is the prototypical example of how variable the spatialization of abstract concepts may be (for overviews, see Galton, 2011; Bender and Beller, 2014). Besides the *linear* representation, which has invited the image of a mental time line, time can also be represented as cyclically recurring (Le Guen and Pool Balam, 2012) or as radially extending from (or pointing toward) one's own present (Bennardo, 2009; Bender et al., 2010). The latter concept in particular, with its half-axes radiating out from the conceptual (deictic) center, factually precludes the existence of a single time line. And it is claimed that some groups like the Yucatec Maya or Amondawa do not represent time in terms of space at all (Sinha et al., 2011; Le Guen and Pool Balam, 2012).

But even those representations that are compatible, in principle, with a linear spatial construct may still vary regarding the number of different time lines a person can hold (e.g., Miles et al., 2011); regarding the axis (i.e., lateral, sagittal, or vertical) along which the lines unfold and the direction in which they point (e.g., Fuhrman et al., 2011; Bergen and Chan Lau, 2012); and regarding whether the lines are anchored in the speaker's subjective present or in objective features of, say, the landscape (e.g., Boroditsky and Gaby, 2010; Núñez et al., 2012a). Part of this variability is due to the fact that, also for spatial representations, we do not just have one available option, and our preferences depend on a bunch of partly unrelated factors, including the perspective focused on and the affordances and constraints inherent in the tasks used.

Whether a similar degree of variability may also be found for the MNL has not yet been investigated in a systematic manner, but possible sources and types of variation have been discussed (e.g., Galton, 1883; Ernest, 1986; Bender and Beller, 2011; Núñez, 2011; Winter et al., 2015), and some characteristics of the MNL are known to vary due to cultural influences (see next section). Yet, with the remarkable degree of variability even *within* one domain, it is almost obvious that there cannot be a simple congruence of the spatially grounded, mental representations of time (MTL) and number (MNL) *across* domains either.

Cultural Impact

A great deal of the variability reported in the previous section can be accounted for by cultural influences, including linguistic metaphors, culture-specific concepts, and culturally embedded practices. For instance, not only the choice of a specific conceptualization of time (i.e., as linear, cyclical, or radial), but also the dimension and direction of linear representations are affected by cultural beliefs and epistemological frameworks, implying, for instance, how the future relates to the present, or whether the future is located in front of or behind the speaker (e.g., León-Portilla, 1990; Núñez and Sweetser, 2006). If time is represented along a linear axis, its direction appears to be additionally influenced by linguistic metaphors, as reflected in expressions such as “looking forward to the future” (inviting a sagittal line pointing from back to front) or “a custom

handed down to us from our ancestors” (inviting a vertical line pointing downwards). Moreover, cultural practices such as those underlying the preferential reading and writing direction appear to be correlated with the direction of the time line (e.g., Tversky et al., 1991; Bergen and Chan Lau, 2012).

A similar influence of the reading and writing direction has also been observed for the mental number line, which extends from left to right for speakers of English, but from right to left for speakers of Arabic and Hebrew (Dehaene et al., 1993; Zebian, 2005; Shaki et al., 2009; for an additional or alternative influence of finger counting, see also Fischer and Brugger, 2011; Bender and Beller, 2012). The second feature of the MNL which is subject to cultural influences is its scale: initially logarithmic, it seems to shift toward linearity with the extent of formal mathematical education (Dehaene et al., 2008), even though some interpret the available data as a composition of two distinct number lines rather than the transformation of one into another (e.g., Moeller et al., 2009).

Finally, both for time and for number, representations may not be spatialized at all (overview in Bender and Beller, 2014), at least not along a spatially extended line. As convincingly argued by Núñez (2008, 2011), the number line is actually a highly sophisticated and culturally mediated concept that took centuries to develop in a particular cultural and historic context, strongly linked to cultural practices of measuring and to instruments such as rulers. Once in place, these practices give rise to SNARC-like effects, not only for quantity representations, but for all kinds of sorting tasks, also for non-numerical categories (Núñez et al., 2011). In untrained participants, at least some response patterns are more accurately accounted for by nonlinear representations (Núñez et al., 2012b), and the extent to which they are spatial to begin with partly depends on the task used (see next section). Clearly, there is a dire need for more research into the exact nature of number representations not affected by Western schooling (Beller et al., 2018).

Dependency on Tasks

A third complication in the picture of cross-domain congruency emerges from the observation that response patterns may depend on task specifics and context.

Again, let us begin with the domain of time, for which this has been analyzed in detail. The toolkit of tasks used to investigate spatial representations of time includes different paradigms: language elicitation, observation of co-speech gestures and postural sway, mapping tasks, and reaction time paradigms based on congruency priming. Notably, the observed time line was found to differ profoundly in terms of the axis along which it unfolds, depending on the paradigm used to investigate it. English speakers, for instance, exhibit a sagittal time line in linguistic tasks and when measuring postural sway, but rely almost exclusively on the lateral axis in co-speech gestures and for tasks requiring spatial layouts (for an overview, see Bender and Beller, 2014). Even within a single paradigm, specifically the tasks based on congruency priming, a variety of axes can be activated depending on task-specific characteristics (Torralbo et al., 2006).

Response patterns may also depend on whether the task explicates the issue of mapping or leaves it implicit. Normally, space-time mappings appear to be highly automatized. Co-speech gestures, for instance, are produced spontaneously without people necessarily being aware of them. In such cases, English speakers strongly prefer to recruit the lateral axis, with leftward gestures for earlier times and rightward gestures for later times. If these same people are asked to *deliberately* produce gestures referring to past and future events, however, they do so much more often along the sagittal axis, familiar to them from linguistic metaphors (Casasanto and Jasmin, 2012).

Crucially, apart from the purely linguistic tasks, most of the tasks typically used in this field contain a spatial component: Co-speech gestures and postural sway inevitably unfold in space, and this is also true for abstract pointing and the arrangement of tokens required in mapping tasks and for the predefined congruency priming in reaction time paradigms. It is thus not surprising that these tasks uncover *spatialized* representations of time. The same arguably holds for the domain of number, where both SNARC effect studies and number line estimation tasks *a priori* impute the spatial representation they try to measure (for related arguments, see also Núñez, 2011; Shaki and Fischer, 2018).

In other words: Spatial representations of number—as of time—may be more diverse than we tend to assume; but in order to explore this realm of possibilities, we require tasks that allow participants to recruit other dimensions beyond the well-known number line; we need to pay attention to perspective and frames of reference; and we need to take the diverse sources of linguistic and cultural variability more seriously. The study reported here, while exploratory in nature, is intended as a first step in this direction. It is based on a paper-and-pencil survey conducted among speakers of five languages: English, Norwegian, German, Chinese, and Japanese.

THE STUDY

Given that both time and number appear to be represented in terms of space, the study reported here aims at exploring the extent to which spatial representations of number may be subject to the same degree of variability, cultural impact, and dependency on task specifics and context as are spatial representations of time. To this end, the study focuses on the extent to which spatial representations of symbolic number depend (i) on a particular perspective or frame of reference, (ii) on the linguistic and cultural background of participants, and (iii) on a specific task.

Regarding (i), spatial representations and inferences change fundamentally depending on which perspective is taken, that is, whether a superordinate field, a given reference point, or a subjective viewpoint is taken as the underlying frame of reference (Levinson, 2003; Majid et al., 2004; Haun et al., 2011). Representations also change depending on whether objects are at rest (static) or moving (dynamic), with assignments of FRONT and FORWARD sometimes flipping between tasks that require a token either to be picked or moved (Bender et al., 2012). While similar dependencies have been observed for temporal

representations, little is known about whether they may also be in place for number representations. We therefore collected data for fixed (static) relations vs. changing (dynamic) relations between specified numbers and number sequences, and data on whether a spatial orientation can be assigned to number sequences and the number line itself.

Regarding (ii), as preferences for a specific frame of reference in the domain of space do vary substantially across languages and cultures (Senft, 1997; Majid et al., 2004; Beller et al., 2015; Beller and Bender, 2017), a corresponding variability in number representations should be observed if these are really grounded in spatial representations. We therefore collected data from speakers of several languages that differ not only in writing systems, traditional writing direction, patterns of finger counting, and/or composition of number words (as detailed below), but also in their preferences for spatial frames of reference.

And regarding (iii), while spatial representations of number yielded with number line estimation tasks and in SNARC studies are necessarily confounded with the spatial layout of the tasks themselves, linguistic tasks offer more leeway for participants to provide responses that need not be compatible with a spatial representation. We therefore collected linguistic data in a questionnaire that explicitly asks participants, in different ways, about the orientation of numerical representations.

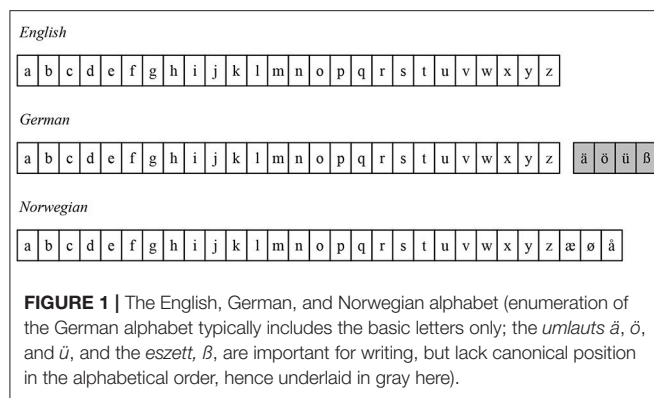
Selection of Languages

The survey was conducted with speakers of five languages: English, Norwegian, German, Chinese, and Japanese. English, Norwegian, and German belong to the Germanic branch of the Indo-European language family. While the two East Asian languages in our sample have no “genetic” relationship, Japanese has been influenced by Chinese in several ways, including with regard to the writing system, parts of the vocabulary, and the number system. The languages were chosen because they differ on several potentially relevant dimensions, including traditional writing direction, preferences for spatial and temporal referencing, and properties of the counting systems.

Writing Systems in the Selection of Languages

The writing systems of the Germanic languages are based on Latin, with a few additional letters in the case of Norwegian and German. The alphabet begins with “a” in all three languages, and ends in “z” in both English and German, and in “å” in Norwegian (Figure 1). All three are written from left to right, with lines ordered from top to bottom.

The standardized form of spoken Chinese is written with logograms (i.e., Chinese characters) in one of two versions: the Simplified Chinese character system prevailing in the People's Republic of China, and the traditional system used outside mainland China. The written standard of Japanese uses mainly two types of writing systems: a set of logograms based on Chinese characters (*kanji*) and two syllabic scripts (*kana*). For the two syllabaries, *hiragana* and *katakana*, the modern and prevalent ordering system *gojuon* is based on 2-dimensional tabulation, beginning with vowel “a” in the upper left corner and ending on “wo” in the bottom right corner (Figures 2A,B). By contrast, *kanji*, with its thousands of symbols, resists conventional



ordering and is therefore sorted according to the composition principles on which the characters are based (for an example, see **Figure 2C**). Traditionally, Chinese and Japanese were written from top to bottom, but writing from left to right is becoming increasingly frequent.

Representations of Space and Time in the Selection of Languages

For describing spatial relations, speakers of all five languages make use of all known basic frames of reference (FoR), but differ with regard to the variant of the relative FoR. These variants differ in how the coordinate system that informs the viewpoint of an observer is transferred to the reference point for localizing an object in relation to the reference point. The object “in front of” the reference point would be the nearer object in the reflection variant, but the further-away object in the translation variant (Levinson, 2003). Germans strongly prefer the reflection type, while the others use both the reflection and translation type, but in distinct proportions (Beller et al., 2015; Beller and Bender, 2017; and see Bender et al., 2012).

With regard to time, speakers of most of these languages appear to recruit at least two distinct axes, and one of these in both directions, depending on task and context (data on English, German, and Chinese are reviewed in Bender and Beller, 2014; for data on Norwegian, see Bender et al., 2017; relevant data on Japanese are still lacking). All recruit the *sagittal* axis with a preference for back-to-front when representing past versus future events, yet with a preference (among German and Chinese speakers) or ambivalence (among English and Norwegian speakers) for the reversed front-to-back when events are moved forward. The *lateral* axis left-to-right is additionally or exclusively used in tasks that recruit space as the medium for representing time, such as in sign language or on paper. Finally, Chinese speakers also recruit the *vertical* axis top-to-bottom.

Representations of Number in the Selection of Languages

The number system in each of the five languages is largely decimal, both for number words and for numerical notations, but composition in the Germanic languages is substantially less regular and transparent than in the Eastern Asian languages (cf., Miura, 1987; Calude and Verkerk, 2016). Notations are

based almost exclusively on Arabic digits for speakers of the Germanic languages, but to some extent also for speakers of Chinese and Japanese, alongside the more traditional Chinese characters. Studies on the mental number line indicate an alignment of the number line primarily with reading and writing direction, in that speakers of English and German exhibit left-to-right orientation (overview in Göbel et al., 2011) and Chinese speakers left-to-right or top-to-bottom orientation, depending on whether numbers are presented as Arabic digits or as Chinese characters (Hung et al., 2008). Japanese speakers, by contrast, were found to respond with left-to-right and bottom-to-top (Ito and Hatta, 2004). Patterns of finger counting which are additionally or alternatively assumed to affect the mental number line (Fischer and Brugger, 2011) are largely similar for speakers of English, Norwegian, and German, but more different for speakers of Japanese, and even more different for speakers of Chinese (overview in Bender and Beller, 2012).

Language-Specific Possibilities for the Spatialization of Number

It may thus be expected that, if number representation follows the preferences for spatial representations, gradual differences between the five languages should be observed, in line with the difference in the degree to which speakers of these languages prefer the reflection versus the translation variant of the relative frame of reference. If number representation follows the preferences for temporal representations, German and Chinese should pattern alike, and should be distinct from English and Norwegian. Finally, if cultural and linguistic factors such as the direction of writing and reading or the transparency of number words play a crucial role, then English, Norwegian, and German should pattern alike, and should be distinct from Chinese and Japanese (which may also differ from one another due to different finger counting patterns).

Methods

Samples

A total of 475 individuals participated in this study. Seven participants were excluded due to being non-native speakers in their respective sample.

The English-speaking sample of 62 individuals was recruited at the University of Nottingham, Great Britain. Most participants were students; 33 (53.2%) were female; and their mean age was $M = 19.8$ years ($SD = 2.6$, range 18–36).

The Norwegian-speaking sample of 78 individuals was recruited at the University of Bergen, Norway. Most participants were students; 59 (75.6%) were female; and their mean age was $M = 25.3$ years ($SD = 7.6$, range 19–62; five did not indicate their age).

The German-speaking sample of 116 individuals was recruited at the University of Freiburg, Germany. Most participants were students; 74 (63.8%) were female; and their mean age was $M = 23.0$ years ($SD = 4.9$, range 18–47; two did not indicate their age).

The Chinese-speaking sample of 89 individuals was recruited from the Chinese community in Freiburg, Germany, and from short-term language courses for foreign students at the University of Freiburg. Most participants were students; 62

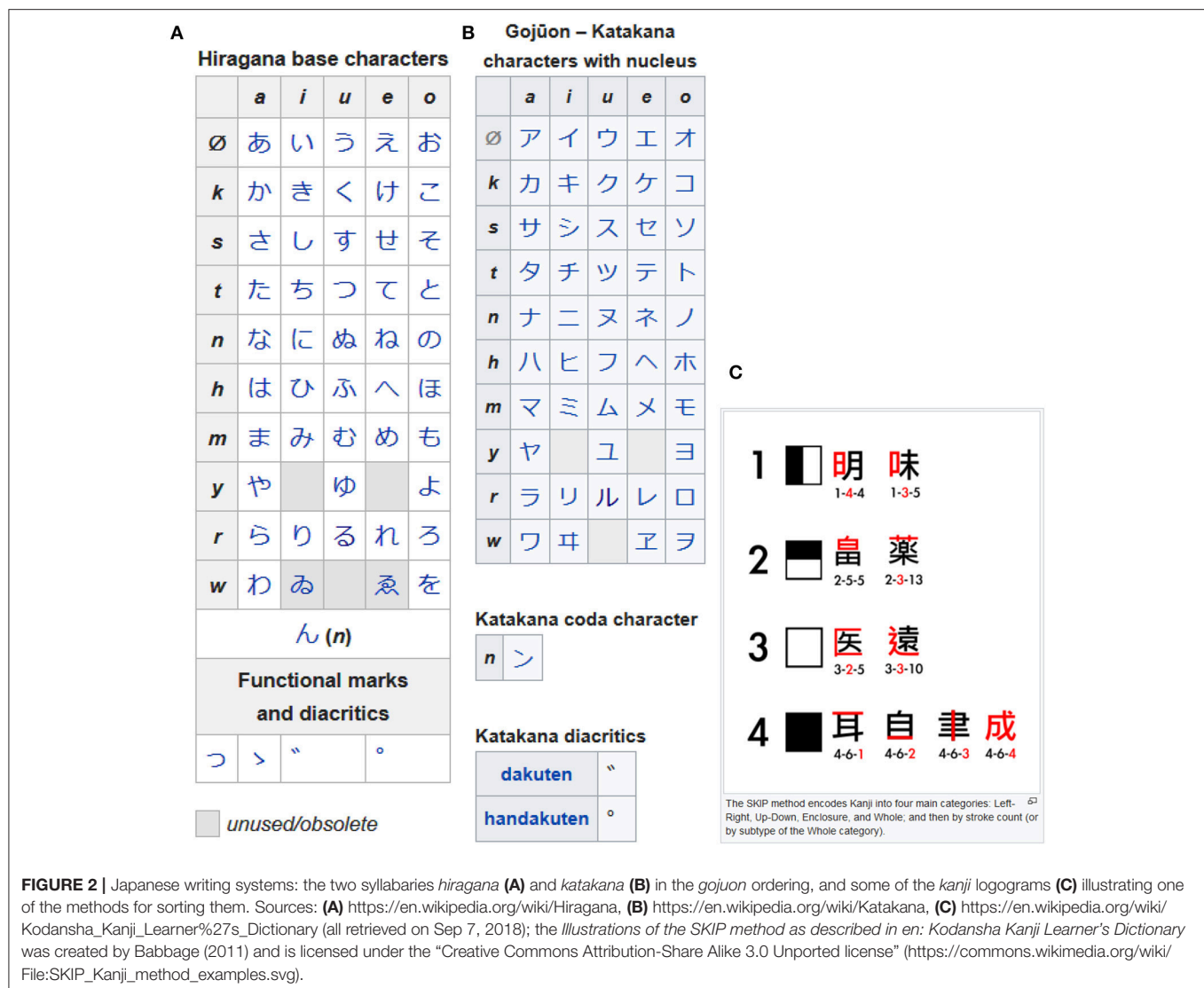


FIGURE 2 | Japanese writing systems: the two syllabaries *hiragana* (A) and *katakana* (B) in the *gojūon* ordering, and some of the *kanji* logograms (C) illustrating one of the methods for sorting them. Sources: (A) <https://en.wikipedia.org/wiki/Hiragana>, (B) <https://en.wikipedia.org/wiki/Katakana>, (C) https://en.wikipedia.org/wiki/Kodansha_Kanji_Learner%27s_Dictionary (all retrieved on Sep 7, 2018); the *Illustrations of the SKIP method as described in en: Kodansha Kanji Learner's Dictionary* was created by Babbage (2011) and is licensed under the “Creative Commons Attribution-Share Alike 3.0 Unported license” (https://commons.wikimedia.org/wiki/File:SKIP_Kanji_method_examples.svg).

(69.7%) were female; and their mean age was $M = 25.5$ years ($SD = 3.4$, range 18–38; four did not indicate their age).

The Japanese-speaking sample of 123 individuals, finally, was recruited at Nagoya University, Japan. All participants were students; 41 (33.6%) were female (one did not indicate his or her gender); and their mean age was $M = 19.5$ years ($SD = 1.8$, range 18–34; two not did not indicate their age).

Materials

The tasks described in the following were part of a larger paper-and-pencil survey that also included temporal and purely spatial items. Here, we focus only on those items that contain numbers or other ordered sequences such as letters or text segments, which are relevant for the questions under scrutiny in this paper. Letters in the tasks on letters ($\alpha 1$ – $\alpha 4$) were based on the Latin alphabet for all but Japanese speakers, for whom the *hiragana* in the *gojūon* ordering was used instead. In the following, we use the British English version for illustration; for translations into Norwegian, German, Mandarin Chinese, and Japanese, see section 1 of

the **Supplementary Materials**. Translations were conducted by bilingual speakers and subsequently back-translated.

(1) The *Moving Task* (Mov) consisted of four items, with an entity to be moved forward or backward (Norwegian: *fram* or *bakover*; German: *nach vorne* or *nach hinten*; Chinese: 往前 or 往后; Japanese: 前に or 後ろに) in the given context. Two items referred to numerical entities:

- (Mov_n1) The 7th signpost was moved {forward/backward} by two positions. Which number does the signpost now have?
- (Mov_n2) Jenny wanted to marry on the 15th of August, but the date had to be moved {forward/backward} by 7 days. On what date does the wedding now take place?

The other two items referred to other ordered entities:

- (Mov_α1) If, in the English alphabet, the letter “E” were moved {forward/backward} by one position, between which two letters would it end up?

(Mov_s1) If, in this sentence, the word “apple” were moved {forward/backward} by three positions, between which two words would it end up?

The items were implemented in four arrangements, crossing, between subjects, two *phrasings* with two *orders of items*. Regarding the phrasings, two items requested a “forward” movement (e.g., *n2* and $\alpha 1$), the other two items a “backward” movement (e.g., *n1* and *s1*), and vice versa¹. One item order was *n2*, *n1*, *s1*, and $\alpha 1$, the other was the exact reversal.

(2) The *Order Task (Ord)* consisted of five items that asked for the order of entities, that is, whether a target entity is in front of or behind (Norwegian: *foran* or *bak*; German: *vor* or *hinter*; Chinese: 前面 or 后面; Japanese: 前 or 後) a reference entity. Two items used a forced-choice format, three used an open format. Two items referred to numerical entities:

(Ord_n3) Number 25 is two positions ...

- ☐ in front of
- ☐ behind
- ... number 23.

(Ord_n4) Which number is 5 positions {in front of/behind} 9?

The other three items referred to other ordered entities:

(Ord_α2) In the alphabet, the letter M is ...

- ☐ in front of
- ☐ behind
- ... the letter P.

(Ord_α3) Which letter is directly {in front of/behind} G in the alphabet?

(Ord_s2) In this sentence, which word is two positions {in front of/behind} the underlined word “two”?

The items were implemented in four arrangements, crossing, between subjects, two *phrasings* with two *orders of items*. The phrasing concerned the order of response options for the forced-choice items (“in front of” first vs. “behind” first) and the preposition used for the open items (“in front of” for the items *n4* and *s2*, and “behind” for the item $\alpha 3$, or vice versa). One item order was determined randomly, the second order was the exact reversal.

(3) The *FRONT Assignment Task (Ass)* consisted of four items that directly asked whether or not an ordered sequence of entities has a front or back (Norwegian: *forside* or *bakside*; German: *Vorne* or *Hinten*; Chinese: 前面 or 后面; Japanese: 前方 or 後方), and if so, in which direction it is pointing. All items followed the same schema and had four response options, exemplified here for the item on the number list:

(Ass_n5) {Front/Back} of an ordered number list ...

- ☐ is at the smallest number.
- ☐ is at the largest number.
- ☐ Something like this does not exist.
- ☐ Something else, namely _____.

As the last two response options were the same for all items, we explicate only the item-specific options for the three remaining items on other ordered sequences:

(Ass_α4) {Front/Back} of the English alphabet ...

- ☐ is at the letter “a”.
- ☐ is at the letter “z”.

(Ass_w) {Front/Back} of the word “holiday” ...

- ☐ is at the letter “h”.
- ☐ is at the letter “y”.

(Ass_q) {Front/Back} of a questionnaire ...

- ☐ is at the instruction part.
- ☐ is at the thanking part.

The items were implemented in four arrangements, crossing, between subjects, two *phrasings* (asking for all items either for “Front of X ...” or “Back of X ...”) with two *orders of items* (one random order and the exact reversal).

Design and Procedure

Four versions of questionnaires were constructed. The various types of tasks were presented within subjects in a fixed order (i.e., the Moving Task followed by the Order Task followed by the FRONT Assignment Task) in line with the increasingly explicit nature of the task (asking for the “front” of an ordered number list highlights the topic of interest more strongly than asking for the date to which an event is moved). The four item arrangements of each task were randomly assigned to one of the four versions of questionnaires, and varied between subjects, as indicated in the Materials section. Participants were instructed to work on all tasks in the given order.

Results

After some preliminaries describing data coding and the procedure for analyzing the single items, we present three types of analyses, separately for numerical and other (i.e., alphabetical and textual) items: item-level analyses, an analysis of participants’ individual consistency across items, and an analysis that helps to decide to what extent the observed variation is task-specific or item-specific.

Preliminaries

Our tasks required participants to indicate a moving direction (Moving Task), a succession (Order Task), or an orientation (FRONT Assignment Task), depending on a specific phrasing (i.e., forward/backward, in front of/behind, and front/back, as described in the Materials section). To enable a comparison of the responses across the phrasings, we re-coded the responses as to whether they indicated that FRONT of the moving direction, of the reference entity, and of the figure’s orientation, respectively, points (i) toward the smallest or largest number of a number sequence for the numerical items, (ii) toward the beginning or the end of the alphabet/hiragana for the alphabetical items, and (iii) toward the beginning or the end of a written segment for the textual items. For answers that did not allow an unambiguous re-coding of FRONT, a missing value was assigned.

¹By accident, for the Japanese version of the item *Mov_n2*, only the forward movement was implemented in all versions of the questionnaire.

We began the data analysis by testing the responses, for each item separately (*item-level analyses*), for differences between languages, and for possible effects of the phrasings and item orders. To this end, we ran a log-linear analysis (Kennedy, 1992) on the re-coded responses with three independent variables: *language* (five versions), *phrasing* (two versions), and *order* of items (two versions). Main effects and interactions were tested for significance by comparing two log-linear models that differ in one candidate factor only.

The analysis started with the full model including the main effects and interactions of all factors. Then, we simplified the model stepwise by excluding one candidate factor at a time (as the basis for the next comparison) in the following order: (1) *language* \times *phrasing* \times *order*, (2) *phrasing* \times *order*, (3) *language* \times *order*, (4) *order*, (5) *language* \times *phrasing*, (6) *phrasing*, (7) *language*. We began with those candidate factors that include the order of items, as we did not expect to find effects of this control variable, and then inspected effects of phrasing and language. Fit values of the computed log-linear models and significance values of the various main effects and interactions are reported in section 2 of the **Supplementary Materials** for each item.

FRONT Assignment on Numerical Items

In the following, we first describe the results of the item-level analyses. Then, we determine participants' individual consistency in assigning FRONT, and inspect possible sources of the observed variation.

(1) *Item-level analyses*. The log-linear analyses indicated a strong main effect *language* for each of the five numerical items, and a modulating effect of how the items were phrased for four items. Participants' FRONT assignments depending on language and phrasing are reported in **Table 1**.

For the item *Mov_n1*, the main effect *language* ($G^2[4] = 88.941$; $p < 0.001$) was the only significant effect. Assignment of FRONT to the smallest number was frequent among speakers of German (86.7%), Chinese (93.2%), and Japanese (85.2%), and less frequent among speakers of English (52.5%) and Norwegian (42.3%).

For the item *Mov_n2*, the analysis revealed two significant effects: a main effect *language* ($G^2[4] = 154.410$; $p < 0.001$) and a small but significant three-way interaction *language* \times *phrasing* \times *order* ($G^2[3] = 9.969$; $p = 0.019$). Again, assignment of FRONT to the smallest number was frequent among speakers of German (94.0%), Chinese (98.9%), and Japanese (98.3%), and less frequent among speakers of English (61.3%) and Norwegian (41.0%). The interaction indicated minor moderating effects of the phrasing and item order.

For the item *Ord_n3*, the analysis revealed two significant effects: a main effect *language* ($G^2[4] = 61.638$; $p < 0.001$) and a small but significant interaction *language* \times *phrasing* ($G^2[4] = 10.429$; $p = 0.034$). Assignment of FRONT to the smallest number was frequent among speakers of German (94.0%), Chinese (98.9%), Japanese (81.3%), and, this time, also Norwegian (78.2%), and less frequent among speakers of English (56.7%). The interaction reflected differences between

TABLE 1 | Percentage of (*N*) participants assigning FRONT to the smallest number of a sequence for the five numerical items (n1 to n5), depending on language and phrasing.

		English		Norwegian		German		Chinese		Japanese	
		Phrasing		Phrasing		Phrasing		Phrasing		Phrasing	
		Forward	Backward	Forward	Backward	Forward	Backward	Forward	Backward	Forward	Backward
MOVING TASK											
Mov_n1	FRONT=smallest ^a	53.3 (30)	51.6 (31)	47.4 (38)	37.5 (40)	86.2 (58)	87.3 (55)	90.9 (44)	95.5 (44)	95.1 (61)	75.4 (61)
Mov_n2	FRONT=smallest ^a	67.7 (31)	54.8 (31)	40.0 (40)	42.1 (38)	89.7 (58)	98.3 (58)	100 (44)	97.7 (44)	98.3 (121)	– (–)
		Phrasing		Phrasing		Phrasing		Phrasing		Phrasing	
		In front of	Behind	In front of	Behind	In front of	Behind	In front of	Behind	In front of	Behind
ORDER TASK											
Ord_n3	FRONT=smallest ^a	71.0 (31)	41.4 (29)	70.0 (40)	86.8 (38)	96.6 (58)	91.4 (58)	100 (44)	97.8 (45)	83.6 (61)	79.0 (62)
Ord_n4	FRONT=smallest ^a	80.0 (30)	56.3 (32)	87.5 (40)	59.5 (37)	100 (58)	86.2 (58)	100 (45)	100 (44)	100 (61)	78.7 (61)
		Phrasing		Phrasing		Phrasing		Phrasing		Phrasing	
		FRONT	BACK	FRONT	BACK	FRONT	BACK	FRONT	BACK	FRONT	BACK
FRONT ASSIGNMENT TASK											
Ass_n5	FRONT=smallest	90.0 (30)	31.3 (32)	61.0 (41)	19.4 (36)	89.7 (58)	77.6 (58)	84.1 (44)	77.3 (44)	93.5 (62)	90.2 (61)
	FRONT=largest	0.0	15.0	0.0	5.6	3.4	12.1	2.3	0.0	1.6	0.0
	Does not exist	10.0	43.7	34.1	63.9	3.4	8.6	6.8	15.9	4.8	8.2
	Other	0.0	0.0	4.9	11.1	3.4	1.7	6.8	6.8	0.0	1.6

^a Percentage FRONT=largest is 100 – percentage FRONT=smallest.

the two phrasings, mainly for English and Norwegian. For English, assignment of FRONT to the smallest number was more frequent when the item asked whether a number is “in front of” another number (71.0%) than when it asked whether a number is “behind” another number (41.4%). The pattern was reversed for Norwegian: Assignment of FRONT to the smallest number was less frequent when the item was phrased with “in front of” (70.0%) than when it was phrased with “behind” (86.8%). For the other languages, the difference between the two phrasings was only marginal ($\leq 5.2\%$).

For the item *Ord_n4*, the analysis again revealed two significant main effects: *language* ($G^2[4] = 54.526$; $p < 0.001$) and *phrasing* ($G^2[1] = 34.609$; $p < 0.001$). This time, FRONT was preferably assigned to the smallest number in all languages: highly frequent among speakers of German (93.1%), Chinese (100%), and Japanese (89.3%), and less so, but still frequent, among speakers of English (67.7%) and Norwegian (74.0%). Overall, this preference varied with the phrasing of the item: It was stronger when the item asked whether a number is “in front of” another number (95.3%) than when it asked whether a number is “behind” another number (78.4%).

Finally, for the item *Ass_n5*, the analysis revealed three significant effects: two main effects, *language* ($G^2[12] = 109.000$; $p < 0.001$) and *phrasing* ($G^2[3] = 32.832$; $p < 0.001$), and an interaction *language* \times *phrasing* ($G^2[12] = 21.732$; $p = 0.041$). Again, assignment of FRONT to the smallest number was frequent among speakers of German (83.6%), Chinese (80.7%), and Japanese (91.9%), and less frequent among speakers of English (59.7%) and Norwegian (41.6%). Overall, this preference was stronger when the item asked participants to indicate the “front” of an ordered number list (smallest: 84.7%; largest: 1.7%; does not exist: 10.6%; other: 3.0%) than when it asked them to indicate the “back” of such a list (smallest: 65.4%; largest: 7.4%; does not exist: 23.4%; other: 3.9%). For a substantial proportion of participants, an ordered number list apparently lacks a front or back. This response was particularly frequent among the English- and Norwegian-speaking participants (27.4 and 48.1%, respectively), as indicated by the significant interaction.

On the whole, the data of the numerical items revealed a quite uniform assignment of FRONT to the smallest number for German, Chinese, and Japanese, and more mixed assignments of FRONT for English and Norwegian. As expected, the control variable *item order* had little influence. The different phrasings played a role in four of the five items, suggesting that the assignment of FRONT to the smallest number was more pronounced when an item asked whether something is “in front of” or is the “front” of a reference entity, but the pattern was not completely homogeneous. Regarding the three types of tasks, the results were fairly homogeneous in all samples except for the Norwegian one; there, the modal response switched from an assignment of FRONT to the largest number in the Moving Task to an assignment of FRONT to the smallest number in the Order Task, and to “Something like that does not exist” in the FRONT Assignment Task.

So far, we have inspected each item separately. In the following, we determine the extent of variation across items by looking at participants’ individual consistency, and we

determine possible sources of the observed variation by looking at participants’ individual response patterns.

(2) *Individual consistency*. In order to obtain an overall measure that reflects the extent to which a participant’s responses vary across items, we counted how often FRONT was assigned to the smallest number and how often it was assigned to the largest number, respectively, across the N numerical items that a participant had solved². For example, if FRONT was assigned to the smallest number on five out of the $N = 5$ items, consistency would be 100% for “FRONT = smallest”; if FRONT was assigned to the smallest number on three items, to the largest number on one item, and was claimed to be “non-existent” on the final item (*Ass_n5*), consistency would be 60% for “FRONT = smallest” and 20% for “FRONT = largest”; and if FRONT was assigned to the smallest number on three items and to the largest number on one item out of $N = 4$ items (one missing response), consistency would be 75% for “FRONT = smallest” and 25% for “FRONT = largest.” We then used the *maximum* of the two counts as an estimate of a participant’s consistency across the whole set of items (i.e., 100, 60, and 75% respectively in the examples).

Across the five samples, FRONT was assigned to either the smallest or the largest number with a mean consistency of 85.3%. An analysis of variance indicated significant differences between the languages; $F_{(4, 463)} = 55.212$; $p < 0.001$; $\eta^2 = 0.323$. Consistency across the five numerical items was high for the speakers of German (91.7%), Chinese (94.5%), and Japanese (89.4%), and was lower for the speakers of English (71.7%) and Norwegian (69.9%). *Post-hoc* tests (Bonferroni-corrected for multiple comparisons) revealed that English and Norwegian did not differ from one another ($p = 1.0$), but both differed from each of the other three languages ($p < 0.001$), and that German, Chinese, and Japanese did not differ from one another ($p > 0.103$).

The consistency values indicate that in general, the individual participant responded in a quite uniform manner, but these values also leave room for variation across samples (particularly for English and Norwegian), across the different types of tasks (Mov, Ord, vs. Ass), and across the adopted FRONT assignment (to the smallest vs. the largest number). In the final step, we therefore inspected individual response patterns in order to qualify this variation. Do the responses attest to uniform, task-specific, or item-specific FRONT assignments?

(3) *Individual response patterns*. This analysis was restricted to those participants who solved all five numerical items. First, we identified participants with a *uniform* FRONT assignment either to the smallest or the largest number of a sequence across the five items. The remaining participants were then checked for *task-specific* response patterns. We determined whether or not the two items of the Moving Task were solved uniformly and whether or not the two items of the Order Task were solved uniformly, by assigning FRONT either to the smallest or to the largest number. Cases with inconsistent FRONT assignments constitute *item-specific* response patterns. The FRONT Assignment Task was not considered here as it consists of only one item of this

² N equals five numerical items minus the number of missing responses.

TABLE 2 | Individual response patterns across the five numerical items (in %, with respective *N* given in brackets).

	English (<i>N</i> = 59)	Norwegian (<i>N</i> = 76)	German (<i>N</i> = 113)	Chinese (<i>N</i> = 87)	Japanese (<i>N</i> = 120)
UNIFORM FRONT ASSIGNMENT ACROSS THE THREE TYPES OF TASKS					
FRONT=smallest	15.3 (9)	11.8 (9)	68.1 (77)	75.9 (66)	56.7 (68)
FRONT=largest	1.7 (1)	—	—	—	—
TASK-SPECIFIC UNIFORM FRONT ASSIGNMENTS					
Moving Task					
FRONT=smallest	28.8 (17)	14.5 (11)	16.8 (19)	17.2 (15)	27.5 (33)
FRONT=largest	27.1 (16)	44.7 (34)	3.5 (4)	1.1 (1)	0.8 (1)
<i>Not uniform</i>	27.1 (16)	28.9 (22)	11.5 (13)	5.7 (5)	15.0 (18)
Order Task					
FRONT=smallest	28.8 (17)	55.3 (42)	20.4 (23)	24.1 (21)	15.8 (19)
FRONT=largest	16.9 (10)	15.8 (12)	1.8 (2)	—	1.7 (2)
<i>Not uniform</i>	37.3 (22)	17.1 (13)	9.7 (11)	—	25.8 (31)
<i>M_{task-specific FRONT}</i>	50.8	65.1	21.3	21.3	22.9
<i>M_{item-specific FRONT}</i>	32.2	23.0	10.6	2.9	20.4

Item-specific response patterns are set in italics.

type and hence precludes a distinction between task-specific and item-specific responses. The results are presented in **Table 2**.

In line with the results from the item-level and consistency analyses, a strong difference emerged between German, Chinese, and Japanese on the one hand, and English and Norwegian on the other.

The majority of the German, Chinese, and Japanese participants assigned FRONT uniformly across all items, and always to the smallest number (ranging from 56.7% for Japanese to 75.9% for Chinese). The mean proportion of task-specific FRONT assignments was about 20%, with FRONT assigned to the smallest number being the most frequent task-specific response. This finding indicates that task-specificity in this case does not result from differences in responses, but rather from uniform responses in one task combined with item-specific responses in the other. Finally, the mean proportion of item-specific responses was relatively low (ranging from 2.9% for Chinese to 20.4% for Japanese).

For English and Norwegian, the patterns are quite different. Uniform FRONT assignments across all items were rather infrequent (17.0% for English; 11.8% for Norwegian); in all cases except one, FRONT was again assigned to the smallest number. Instead, the mean proportion of task-specific FRONT assignments was high (50.8% for English; 65.1% for Norwegian). Compared to the other three languages, FRONT was more often assigned to the largest number; in fact, this was the modal response for the Moving Task in the Norwegian sample. Finally, the mean proportion of item-specific responses was higher for English and Norwegian (32.2 and 23.0%, respectively) than for the other three languages.

FRONT Assignment on Alphabetical and Textual Items

As for the numerical items, we begin with the item-level analyses (first for the alphabetical items, and then for the textual items), before determining participants' consistency in assigning FRONT and their individual response patterns across items.

(1) *Item-level analyses.* The log-linear analyses indicated a strong main effect *language* for each of the eight alphabetical and textual items, and a modulating effect of how the items were phrased for six items. Participants' FRONT assignments depending on language and phrasing are reported in **Table 3**.

For the item *Mov_α1*, the analysis revealed two significant effects: a main effect *language* ($G^2[4] = 129.663$; $p < 0.001$) and a small but significant interaction *language* \times *phrasing* ($G^2[4] = 15.765$; $p = 0.003$). Assignment of FRONT to the beginning of the alphabet was frequent among speakers of German (83.2%), Chinese (97.8%), and Japanese (94.2%), and substantially less frequent among speakers of English (45.9%) and Norwegian (41.6%). The interaction reflected differences between the two phrasings, mainly for English and Japanese. For English, assignment of FRONT to the beginning of the alphabet was more frequent when the item asked about a letter being moved "forward" (56.7%) than when it asked about a letter being moved "backward" (35.5%). The pattern was reversed for Japanese: Assignment of FRONT to the beginning of the alphabet was less frequent when the item was phrased with "forward" (87.2%) than when it was phrased with "backward" (100%). For the other languages, the difference between the two phrasings was only marginal ($\leq 6.2\%$).

For the item *Ord_α2*, the main effect *language* ($G^2[4] = 51.692$; $p < 0.001$) was the only significant effect. Different from the item *Mov_α1*, FRONT was preferably assigned to the beginning of the alphabet in all languages: highly frequently among speakers of German (98.3%), Chinese (100%), Japanese (87.0%), and Norwegian (85.9%), and less so but still frequently among speakers of English (71.0%).

For the item *Ord_α3*, the analysis revealed two significant main effects: *language* ($G^2[4] = 68.264$; $p < 0.001$) and *phrasing* ($G^2[1] = 8.227$; $p = 0.004$). Again, FRONT was preferably assigned to the beginning of the alphabet in all languages: highly frequently among speakers of German (99.1%), Chinese (100%), Japanese (96.7%), and Norwegian (88.5%), and less so

TABLE 3 | Percentage of (*N*) participants assigning FRONT to the beginning of a sequence for the four alphabetical items ($\alpha 1$ – $\alpha 4$) and four textual items (s1, s2, w, q), depending on language and phrasing.

		English		Norwegian		German		Chinese		Japanese	
		Phrasing		Phrasing		Phrasing		Phrasing		Phrasing	
		Forward	Backward	Forward	Backward	Forward	Backward	Forward	Backward	Forward	Backward
MOVING TASK											
Mov_α1	FRONT=beginning ^{a,b}	56.7 (30)	35.5 (31)	41.0 (39)	42.1 (38)	80.0 (55)	86.2 (58)	95.5 (44)	100 (45)	87.2 (47)	100 (57)
Mov_s1	FRONT=beginning ^{a,c}	22.6 (31)	35.5 (31)	28.9 (38)	23.1 (39)	84.2 (57)	89.7 (58)	97.8 (45)	93.2 (44)	98.0 (51)	84.8 (46)
		Phrasing		Phrasing		Phrasing		Phrasing		Phrasing	
		In front of	Behind	In front of	Behind	In front of	Behind	In front of	Behind	In front of	Behind
ORDER TASK											
Ord_α2	FRONT=beginning ^{a,b}	80.6 (31)	61.3 (31)	87.5 (40)	84.2 (38)	100 (58)	96.6 (58)	100 (44)	100 (45)	82.0 (61)	91.9 (62)
Ord_α3	FRONT=beginning ^{a,b}	80.0 (30)	50.0 (30)	94.6 (37)	82.9 (41)	98.3 (58)	100 (58)	100 (44)	100 (45)	98.4 (61)	95.2 (62)
Ord_s2	FRONT=beginning ^{a,c}	83.3 (30)	61.3 (31)	93.1 (29)	61.3 (31)	100 (54)	96.6 (58)	97.8 (45)	97.7 (44)	75.0 (4) ^f	92.3 (13) ^f
		Phrasing		Phrasing		Phrasing		Phrasing		Phrasing	
		FRONT	BACK	FRONT	BACK	FRONT	BACK	FRONT	BACK	FRONT	BACK
FRONT ASSIGNMENT TASK											
Ass_α4	FRONT=beginning ^b	93.3 (30)	59.4 (32)	68.3 (41)	30.6 (36)	98.3 (58)	93.1 (58)	100 (45)	97.7 (44)	88.7 (62)	90.2 (61)
	FRONT=end ^b	0.0	15.6	0.0	2.8	0.0	1.7	0.0	2.3	3.2	0.0
	Does not exist	6.7	25.0	29.3	61.1	1.7	5.2	0.0	0.0	8.1	4.9
	Other	0.0	0.0	2.4	5.6	0.0	0.0	0.0	0.0	0.0	4.9
Ass_w	FRONT=beginning ^d	90.0 (30)	50.0 (32)	56.1 (41)	18.9 (37)	98.3 (58)	93.1 (58)	95.6 (45)	97.7 (44)	75.8 (62)	86.9 (61)
	FRONT=end ^d	0.0	18.8	0.0	16.2	0.0	0.0	0.0	0.0	4.8	0.0
	Does not exist	10.0	31.2	43.9	40.5	1.7	5.2	4.4	0.0	14.5	13.1
	Other	0.0	0.0	0.0	24.3	0.0	1.7	0.0	2.3	4.8	0.0
Ass_q	FRONT=beginning ^e	96.7 (30)	78.1 (32)	95.0 (40)	73.0 (37)	93.1 (58)	94.8 (58)	97.8 (45)	100 (44)	74.2 (62)	73.8 (61)
	FRONT=end ^e	0.0	9.4	5.0	5.4	0.0	1.7	0.0	0.0	1.6	3.3
	Does not exist	3.3	9.4	0.0	0.0	1.7	0.0	0.0	0.0	19.4	19.7
	Other	0.0	3.1	0.0	21.6	5.2	3.4	2.2	0.0	4.8	3.3

^aPercentage FRONT=end is 100 – percentage FRONT=beginning.

^bBeginning: first letter of the alphabet ("A"); end: last letter ("Z").

^cBeginning: first word of a sentence; end: last word (read from left to right).

^dBeginning: first letter of a word; end: last letter (read from left to right).

^eBeginning: introduction part of a questionnaire; end: thanking part.

^fThe responses of many Japanese participants could not be coded properly due to an ambiguity in the Japanese version of this item.

but still frequently among speakers of English (65.0%). Overall, this preference was stronger when the item asked participants to indicate whether a letter is “in front of” another letter (95.7%) compared to whether a letter is “behind” another letter (89.4%).

For the item *Ass_α4*, the analysis again revealed two significant main effects: *language* ($G^2[12] = 110.202$; $p < 0.001$) and *phrasing* ($G^2[3] = 18.120$; $p < 0.001$). Assignment of FRONT to the beginning of the alphabet was highly frequent among speakers of German (95.7%), Chinese (98.9%), and Japanese (89.4%), less so but still frequent among speakers of English (75.8%), and least frequent among speakers of Norwegian (50.6%). Overall, this preference was stronger when the item asked participants to indicate the FRONT of the alphabet (FRONT=beginning: 90.3% FRONT=end: 0.8%; does

not exist: 8.5%; other: 0.4%) than when it asked participants to indicate the BACK of the alphabet (FRONT=beginning: 78.8%; FRONT=end: 3.5%; does not exist: 15.6%; other: 2.2%). As with number lists (cf. item *Ass_n5*), for some participants, the alphabet lacks a front or back. This response was given by some English-speaking participants (16.1%) and was particularly frequent among the Norwegian-speaking participants (44.2%).

For the item *Mov_s1*, the analysis revealed two significant effects: a main effect *language* ($G^2[4] = 190.755$; $p < 0.001$) and a small but significant three-way interaction *language* × *phrasing* × *order* ($G^2[4] = 11.247$; $p = 0.024$). Assignment of FRONT to the beginning of a sentence was highly frequent among speakers of German (87.0%), Chinese (95.5%), and Japanese (91.8%), but rather infrequent among speakers of English (29.0%) and

Norwegian (26.0%). The interaction indicated minor moderating effects of the phrasing and item order.

For the item *Ord_s2*, the analysis revealed two significant main effects: *language* ($G^2[4] = 44.824$; $p < 0.001$) and *phrasing* ($G^2[1] = 10.934$; $p < 0.001$). Different from the item *Mov_s1*, FRONT was preferably assigned to the beginning of a sentence in all languages: highly frequently among speakers of German (98.2%), Chinese (97.8%), and Japanese (88.2%), and less so but still frequently among speakers of English (72.1%) and Norwegian (76.7%). Overall, this preference was stronger when the item asked participants to indicate whether a word is “in front of” another word (94.4%) compared to whether a word is “behind” another word (84.2%).

For the item *Ass_w*, the analysis revealed three significant effects: a main effect *language* ($G^2[12] = 126.132$; $p < 0.001$), a main effect *phrasing* ($G^2[3] = 15.471$; $p = 0.001$), and an interaction *language* \times *phrasing* ($G^2[12] = 44.849$; $p < 0.001$). Assignment of FRONT to the beginning of a word was highly frequent among speakers of German (95.7%) and Chinese (96.6%), less so but still frequent among speakers of Japanese (81.3%) and English (69.4%), and least frequent among speakers of Norwegian (38.5%). Overall, this preference was stronger when the item asked participants to indicate the FRONT of a word (FRONT=beginning: 83.5%; FRONT=end: 1.3%; does not exist: 14.0%; other: 1.3%) compared to the BACK of the word (FRONT=beginning: 74.6%; FRONT=end: 5.2%; does not exist: 15.5%; other: 4.7%), but this difference does not hold uniformly for all samples (in fact, it was reversed for Japanese), as indicated by the interaction. For some participants, words lack a front or back. This response was given by some speakers of Japanese (13.8%) and English (21.0%), and was particularly frequent among speakers of Norwegian (42.3%).

Finally, for the item *Ass_q*, the main effect *language* ($G^2[12] = 73.846$; $p < 0.001$) was the only significant effect. FRONT was preferably assigned to the beginning of a questionnaire in all languages: highly frequently among speakers of German (94.0%), Chinese (98.9%), English (87.1%), and Norwegian (84.4%), and less so but still frequently among speakers of Japanese (74.0%). Different from all other items of the FRONT Assignment Task, the response option “Something like front/back does not exist” did not play a major role for most samples ($\leq 6.5\%$), except for the Japanese speakers (19.5%).

On the whole, the data of the alphabetical and textual items revealed a quite uniform assignment of FRONT to the beginning of the alphabet or text segment for German, Chinese, and Japanese, and more mixed assignments of FRONT for English and Norwegian. As expected, the control variable *item order* did not have much of an influence. The different phrasings played a role for six of the eight items, suggesting that the assignment of FRONT to the beginning of the alphabet or text segment was more pronounced when an item asked participants to indicate whether something is “in front of” or is the “front” of a reference entity, but the pattern was not completely homogeneous. Regarding the three types of tasks, the results were fairly homogeneous for German, Chinese, and Japanese, but not for English and Norwegian; there, the modal response switched from an assignment of FRONT to the end of the alphabet or

sentence in the Moving Task to an assignment of FRONT to the beginning in the Order Task, and to an assignment of FRONT to the beginning or to “Something like that does not exist” in the FRONT Assignment Task.

(2) *Individual consistency*. Consistency values across the eight alphabetical and textual items were calculated as described for the numerical items. Across the five samples, FRONT was assigned to either the smallest or the largest number with a mean consistency of 85.0%. An analysis of variance indicated significant differences between the languages; $F_{(4,463)} = 89.087$; $p < 0.001$; $\eta^2 = 0.435$. Consistency was high for the speakers of German (94.4%), Chinese (98.2%), and Japanese (87.3%), and was lower for the speakers of English (69.7%) and Norwegian (64.5%). *Post-hoc* tests (Bonferroni-corrected for multiple comparisons) revealed that English and Norwegian did not differ from one another ($p = 0.354$), but both differed from each of the other three languages ($p < 0.001$); that German and Chinese did not differ from one another ($p = 0.597$), but both differed from each of the other three languages ($p < 0.002$); and that Japanese differed from all other languages ($p < 0.002$).

(3) *Individual response patterns*. As for the numerical items, this analysis was restricted to those participants who solved the relevant alphabetical and textual items, comprising seven items for Japanese (excluding the item *Ord_s2* that could not be coded appropriately for most participants) and all eight items otherwise. First, we identified participants with a *uniform* FRONT assignment either to the beginning or the end of the alphabet or text segment across the whole set of items. The remaining participants were then checked for *task-specific* response patterns. We determined whether or not each of the three types of tasks was solved uniformly, by assigning FRONT either to the beginning or to the end: the Moving Task with two items, the Order Task with three items (Japanese: two items), and the FRONT Assignment Task with three items. Again, cases with inconsistent FRONT assignments constitute *item-specific* response patterns. The results are presented in **Table 4**.

In line with the previous results, a strong difference again emerged between German, Chinese, and Japanese on the one hand, and English and Norwegian on the other.

The majority of the German, Chinese, and Japanese participants assigned FRONT uniformly across all items, and always to the beginning of the alphabet or text segment (ranging from 52.3% for Japanese to 85.4% for Chinese). The mean proportion of task-specific FRONT assignments ranges from 9.7% for Chinese to 30.7% for Japanese, with FRONT assigned to the beginning being the most frequent task-specific response. This finding again indicates that task-specificity in this case does not result from differences in responses, but rather from uniform responses in one task combined with item-specific responses in the other. Finally, the mean proportion of item-specific responses was relatively low (ranging from 4.9% for Chinese to 17.0% for Japanese).

For English and Norwegian, the patterns are quite different. Uniform FRONT assignments across all items were again rather infrequent (11.9% for English and 12.1% for Norwegian); in all cases, FRONT was again assigned to the beginning of the alphabet and text segment. The mean proportion of task-specific FRONT

TABLE 4 | Individual response patterns across the eight (seven for Japanese^a) alphabetical and textual items (in %, with respective *N* given in brackets).

	English (<i>N</i> = 59)	Norwegian (<i>N</i> = 58)	German (<i>N</i> = 109)	Chinese (<i>N</i> = 89)	Japanese (<i>N</i> = 88)
UNIFORM FRONT ASSIGNMENT ACROSS THE THREE TYPES OF TASKS					
FRONT=beginning	11.9 (7)	12.1 (7)	73.4 (80)	85.4 (76)	52.3 (46)
FRONT=end	—	—	—	—	—
TASK-SPECIFIC UNIFORM FRONT ASSIGNMENT					
Moving Task					
FRONT=beginning	13.6 (8)	12.1 (7)	3.7 (4)	7.9 (7)	35.2 (31)
FRONT=end	50.8 (30)	48.3 (28)	7.3 (8)	—	2.3 (2)
<i>Not uniform</i>	23.7 (14)	27.6 (16)	15.6 (17)	6.7 (6)	10.2 (9)
Order Task					
FRONT=beginning	32.2 (19)	53.4 (31)	23.9 (26)	12.4 (11)	36.4 (32)
FRONT=end	8.5 (5)	1.7 (1)	—	—	—
<i>Not uniform</i>	47.5 (28)	32.8 (19)	2.8 (3)	2.4 (2)	11.4 (10)
FRONT Assignment Task					
FRONT=beginning	54.2 (32)	15.5 (9)	19.3 (21)	9.0 (8)	14.8 (13)
FRONT=end	1.7 (1)	—	—	—	—
Does not exist	3.4 (2)	—	0.9 (1)	—	3.4 (3)
Other	—	—	—	—	—
<i>Not uniform</i>	28.8 (17)	72.4 (42)	6.4 (7)	5.6 (5)	29.5 (26)
<i>M_{task-specific FRONT}</i>	54.8	43.7	18.3	9.7	30.7
<i>M_{task-specific FRONT}</i>	33.3	44.3	8.3	4.9	17.0

Item-specific response patterns are set in italics.

^aIn Japanese, the analysis is based on seven items only; the item *Ord_s2* was excluded, because it was solved appropriately only by a handful of participants.

assignments was high (54.8% for English; 43.7% for Norwegian). Compared to the other three languages, FRONT was more often assigned to the end of the alphabet and text segment; in fact, this was the modal response for the Moving Task both in the English and the Norwegian sample. Finally, the mean proportion of item-specific responses was higher for English and Norwegian (33.3 and 44.3%, respectively) than for the other three languages. The high proportion of item-specific responses for Norwegian was mainly due to the FRONT Assignment Task, which showed a particularly high value (72.4%).

DISCUSSION

The main goal of the current study was to explore the potential for variability in spatial representations of number. Specifically, it aimed at investigating the extent to which such representations depend (i) on the perspective taken and other specifics of the tasks, (ii) on the linguistic and cultural background of participants, and (iii) on the research paradigm. While our findings so far paint a rather complex picture, they suggest that the spatial alignment of number representations is indeed more variable than previously assumed, and that all of the factors investigated do affect the alignment. In the following, we first outline and discuss (in reverse order) the emerging patterns for each factor in the numerical tasks, before comparing respective patterns across domains, both with the alphabetical and textual tasks reported above and with similar sets of tasks in the temporal and spatial domain as reported elsewhere.

Sources of Within-Domain Variability in the Numerical Task

Possible sources of the variability in numerical tasks include the research paradigm, cultural and linguistic differences, as well as the perspective chosen and other task specifics.

Research Paradigm

Number representations, if spatialized in a linear manner, may unfold along three distinct dimensions: lateral (i.e., left/right), sagittal (back/front), or vertical (bottom/top), in either direction. Whereas standard paradigms for number line assessment predefine a particular spatial dimension as part of the task (e.g., SNARC tasks typically recruit the lateral dimension), and hence obtain spatial representations along this dimension, the current study used a language-based paradigm to probe whether a different (i.e., the sagittal) dimension may also be recruited for alignment. The findings from the current study indicate that this is indeed the case.

Specifically, numbers may be aligned not only with the *lateral* axis in either direction, as for speakers of the Germanic languages (Dehaene et al., 1993; Siegler and Opfer, 2003; Wood et al., 2008; Moeller et al., 2009), or with the *vertical* axis, as for speakers of Chinese and Japanese (Ito and Hatta, 2004; Hung et al., 2008), but also along the *sagittal* axis. Speakers of German, Chinese, and Japanese exhibited a strong preference for representing smaller numbers “in front of” larger numbers. This preference was less consistent, but nevertheless also observed, among the English and Norwegian speakers. Evidently, alignment of the number line with the sagittal axis makes sense for most of our participants

(for evidence on a similar near-to-far alignment (see Santens and Gevers, 2008).

These findings imply not only that the spatial alignment of the number line may be more diverse than previously assumed, but also—and importantly—that people are prepared to adopt more than one such type of spatialized representation depending on the nature of the task context (see also Hung et al., 2008; Fischer et al., 2010; Winter et al., 2015). A possibility not explicitly tested in the current study, but raised by the parallels between number and time representations, is that distinct ways of anchoring (e.g., in the person him/herself or in external reference points) may also affect how the number line is spatialized (cf. Bender and Beller, 2014). This would also necessitate differentiating more strictly the dimensions under scrutiny and paying more diligence to how they are implemented in the experimental design. When, for instance, the sagittal axis (front/back) is conflated with a radial axis (near/far), or vertical representations are measured with tabletop layouts (i.e., along the sagittal/radial axis), findings and their interpretation are unnecessarily obscured (Winter et al., 2015).

Cultural and Linguistic Differences

Whereas most previous studies interested in the potential of cultural influences focused on the direction of reading and writing as the most obvious factor for shaping the MNL (e.g., Dehaene et al., 1993; Zebian, 2005; Shaki et al., 2009; for a more nuanced perspective see Shaki and Fischer, 2008, 2018; Fischer et al., 2010), we investigated whether native language and/or cultural background may also influence the MNL by other means. The findings reported above seem to confirm this, but inferences so far remain speculative.

Specifically, we did find significant cultural differences, but interestingly not along the lines one may have expected. While speakers of German, Chinese, and Japanese—three entirely unrelated languages—exhibited the same strong preference for the same type of representation (FRONT pointing toward the smallest number), speakers of English and Norwegian—two close relatives of German—differed both from German and from each other. Notably, these differences emerged not so much in terms of different preferences for MNL orientation, but rather in an apparent overall lack of clear preferences on the part of English and Norwegian speakers. That is, within these two groups, not even within-cultural consensus was achieved. While the present findings cannot account for this lack of within-cultural consensus in MNL orientation, it is in line with a similar lack of consensus in MTL orientation for the same populations (Rothe-Wulf et al., 2015; Bender et al., 2017)—a pattern we will come back to in the section below in which we compare patterns across domains.

Since speakers of English, Norwegian, and German share almost identical writing systems—in contrast to Chinese and Japanese speakers—writing and reading direction can be excluded as a relevant factor for the differences observed here. The same is true for a possible influence of the counting system, especially in terms of the transparency and regularity in number word construction and of the patterns of finger counting, which were alternatively discussed as prime factors in shaping spatialized number lines (cf., Bender and Beller, 2011, 2018;

Fischer and Brugger, 2011), as these differ substantially between German, Chinese, and Japanese. Which cultural (or other) factors may be responsible for these differences, then, remains unclear.

Perspective and Other Task Specifics

Whereas a widespread assumption holds a homogeneous concept of the MNL as something rather stable and independent of the perspective taken, research on the domains of space and time points to the possibility that representations and inferences may change according to whether a superordinate field, a given reference point, or a subjective viewpoint is taken as the underlying frame of reference, and according to whether static or dynamic relations are at stake. To examine the potential influence of these factors, we therefore collected data on fixed (static) versus changing (dynamic) relations between specified numbers and number sequences, and on whether a spatial orientation can be assigned to number sequences and the number line itself as the superordinate field. For reasons of control, we also varied the polarity of the spatial expression under scrutiny, that is, whether items were phrased using the formulations “front,” “in front of,” and “forward,” or the reversed set “back,” “behind,” and “backward.”

Somewhat unexpectedly, the specific formulation used (i.e., “front/in front of/forward” vs. “back/behind/backward”) had significant effects on response patterns, and for speakers of English almost reversed the trend across the types of tasks. While this apparently inconsistent usage of complementary poles is hard to account for in the context of our study, it is not an unusual observation (e.g., Grabowski and Weiß, 1996; Grabowski and Miller, 2000). Against this background, in the following, we only consider the results of those tasks that were formulated with “front,” “in front of,” or “forward.”

As detailed above, three of our groups held strong preferences regarding the orientation of the number line along the sagittal axis, namely with FRONT pointing toward the smallest number. While their strong and consensual preference does not leave much space for variation, the two remaining groups were sensitive to task specifics, and the patterns observed suggest that the same distinctions as for space and time may also be decisive for how the number line is oriented.

Specifically, the more explicitly the tasks ask for FRONT in these ordered sequences, the more English speakers indicate the smallest number as in FRONT: most strongly when explicitly assigning FRONT to the number line (FRONT Assignment Task), less so when assessing the order of a sequence (Order Task), and least when moving a number forward (Moving Task). A similar pattern emerges for speakers of Norwegian, except that, in the Assignment Task, a substantially greater number of participants (34% as compared to 10% among English speakers) rejects the notion that a number sequence may have a FRONT (perhaps due to an infelicitous translation of “front”; cf. Bender et al., 2017). Still, almost all of those who do consider this notion sensible agree on where FRONT would be: pointing toward the smallest number and the beginning of a sequence (with 0% pointing to the opposite end of the sequence for all task items except the “questionnaire”). More importantly, this general trend of an increase in FRONT assignment to the smaller number with

increasing explicitness leads to a reversal of preferences for the least explicit task (i.e., the Moving Task). Here, speakers of Norwegian actually assigned FRONT more often to the larger numbers.

Comparison of Patterns Across Domains

To investigate cross-domain correspondences, we compare the patterns of the numerical items first with those from the alphabetical and textual items reported above, and then with data on the temporal and spatial domain, obtained in related studies reported elsewhere. When comparing response patterns across domains, we consider all those items as in the same direction that—on a lateral axis—would be regarded as on the same side: that is, for instance, the smallest number, “a” in the Latin alphabet and the Japanese *kana*, the beginning of a piece of text, and—in the domain of time—the past (for speakers of the languages under scrutiny, all these directions would be localized left); for an overview (see **Figure 3**).

Comparison With the Alphabetical and Textual Domain

Across the numerical, alphabetical, and textual domain, similarities in the response patterns are striking (see **Figure 3**). They are almost perfect for the Assignment Task, with the exception of the proportion to which participants chose the “does not exist” option, and for the Order Task. Patterns are also largely replicated for the Moving Task, yet with even lower assignments of FRONT to the earlier items among speakers of English and Norwegian in the textual items as compared to the others. This coincides with a clear preference for the reversed orientation (i.e., later items as FRONT) in the textual items, whereas the alphabetical and numerical items give rise to more ambivalence among speakers of these two languages.

Comparison With the Temporal and Spatial Domain

To compare the data set reported here with data on the temporal and spatial domain, we draw on previously published findings (Bender et al., 2010, 2012, 2017; Beller et al., 2015; Beller and Bender, 2017). As some of these findings comprise partly different language selections, we lack comparable data for some of the languages in some of the tasks. Furthermore, for the spatial domain, two additional conventions need to be specified. First, a task corresponding to the Assignment Task used here is not possible for space as such because space has no beginning and, due to its greater number of dimensions, also has more degrees of freedom for alignment. Second, in order to establish comparable relations for the Order Task and the Moving Task, we pick those spatial items that contain a deictic center (i.e., an observer) as the component conveying orientation. In the spatial domain, such relations define a relative frame of reference in one of several variants (cf., Levinson, 2003); of the two variants relevant here, *reflection* renders the entity nearer to the observer as in front of the other, whereas *translation* renders the further-away entity as in front.

Interestingly, while the available temporal data—both from speakers of Norwegian obtained with the same set of tasks (Bender et al., 2017) and from speakers of English, German, and

Chinese obtained with a different but structurally similar set of tasks (Bender et al., 2010; see also Rothe-Wulf et al., 2015)—closely reflect the numerical data, this does not hold for the spatial data (see **Figure 3**). In fact, the spatial response pattern is the one that most strikingly *differs* from the response patterns in all other domains. Here, the German pattern is closest to the pattern in the other two Germanic languages and is distinctly different from those in Chinese and Japanese³. Speakers of Chinese and Japanese exhibit an assignment pattern in the Order Task that is opposite to that in the numerical domain, albeit with a good deal of variability. And moving an entity forward strongly triggers FRONT assignment to the further-away entity in all investigated languages alike.

An Account of Cross-Domain Similarities and Differences

As detailed in **Figure 3**, similarities across the domains investigated in the current study (i.e., number, alphabet, and text segments) as well as time are not perfect, but are substantial, for speakers of five different languages. Interestingly, of all domains, it is the spatial domain that does not fit the general pattern. What may explain both the convergence in the former and the disparity in the latter?

Ordered Sequences

To illuminate what we hold to be the underlying mechanism, let us first return to the difference between tasks. All of the ordered sequences used here, including time, are conceived of as having a beginning: the Latin alphabet and the Japanese *kana* (according to the *gojuon* ordering) in the letter for “a,” the sequence of number words in 1, text segments in the first word written, and time in the past. At least metaphorically, beginning corresponds to FRONT. This inherent orientation may also serve for ordering two elements within a sequence, localizing the earlier ones in the sequence as closer to its beginning. For instance, the smaller number, being closer to the FRONT of the number sequence, would therefore be regarded as “in front of” the larger number. A dynamic context such as following the path of the sequence may activate a different perspective—one that shifts the assignment of FRONT into the direction of the movement.

On this account, the Assignment Task should evoke an alignment of FRONT with the beginning of the sequence across languages. In the Order Task, and even more so in the Moving Task, this preference may be superimposed to some extent by a preference for the reversed orientation, in that FRONT is now more readily assigned to the direction of movement (as reflected in expressions like “counting forward/backward”). And indeed, in the Assignment Task, in which motion plays no role, the overwhelming majority of participants assign FRONT to the beginning. In the Moving Task, this preference appears to come into conflict with the reversed preference for dynamic settings, which is why consensus is lowest here. Responses in the Order

³Values for the static task were somewhat lower across the board in the study that also investigated dynamic relations (Bender et al., 2012). There, FRONT was assigned to the nearer token by 56.5, 74.6, and 20.3% of the very same speakers of English, German, and Chinese, respectively (indicated in **Figure 3** by the vertical stroke in the respective bars).

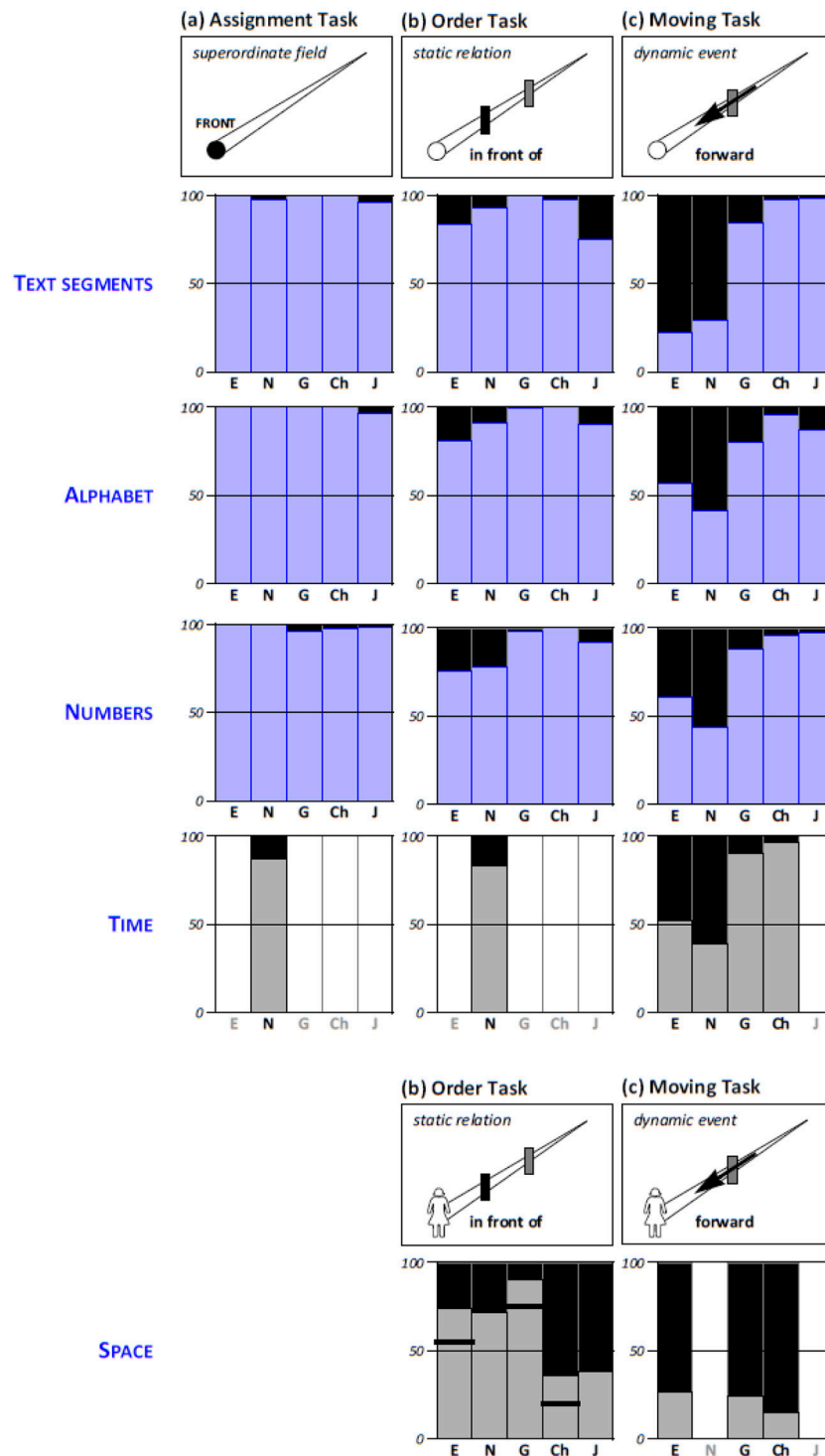


FIGURE 3 | Response patterns across domains. The bars in the panels indicate the proportion of participants (in %) assigning FRONT to the beginning of the ordered sequence (for text segments, the alphabet, number, or time) and toward the observer/Ego (for space). Data for the Assignment Tasks are recalculated to include only those who chose a specific direction; data for the textual and alphabetical items are aggregated across tasks, and data from the spatial domain are aggregated over reflection and rotation (which both imply the nearer item as in front, in contrast to translation). *Sources of additional data:* Time – Norwegian (Bender et al., 2017), English, German, and Chinese (Bender et al., 2010; see also Rothe-Wulf et al., 2015); space/static – English, German, and Chinese (Beller et al., 2015; see also Bender et al., 2012), Norwegian and Japanese (Beller and Bender, 2017); space/dynamic – English, German, and Chinese (Bender et al., 2012). The vertical strokes in the bars for space indicate the somewhat lower values for the static task as collected in the study that also investigated dynamic relations (Bender et al., 2012).

Task are interjacent. These assumptions are compatible with how, across languages and cultures, our participants responded to the tasks; cultural differences mainly emerged with regard to the extent to which the dynamic aspect triggered a reversal of perspectives (for a particularly striking case, see Rothe-Wulf et al., 2015).

This account offers an explanation not only for the convergence across cultures, but also for the convergence across domains. The reason for the latter, we propose, is that all domains—except for space—share important characteristics and may even be based on overlapping representations. For instance, the number sequence and the alphabet are organized in very similar ways, one arguably patterned on the other. Both are ordered sequences, recited endlessly for memorization in childhood, structurally similar to sentences and, when noted down, constituting a specific genre of text. All of these also share characteristics with time. On the one hand, time is generally organized by numbers, most obvious in how we specify date and time. On the other hand, ordered sequences such as numbers or letters also unfold along the temporal dimension: When enumerating the list of number words, reciting the letters of the alphabet, and writing sentences or larger pieces of text, the same process turns future or further-away entities into past and nearer entities. As we recite the sequence of counting words, for instance, it is the smaller ones that move further away into the past as time passes by.

Some empirical support comes from recent work by Sasanguie and colleagues. While taking an entirely different approach, their work confirms a central role of memory processes in the construction of symbolic number representations. Specifically, their findings point to the associations between numbers stored in long-term memory as a key factor for stable numerical representations and arithmetic competence (Sasanguie et al., 2017), thereby also supporting the critical shift from cardinal to ordinal processing in the development of children's numerical understanding (Sasanguie and Vos, 2018). This crucial role of verbal encoding for a linear spatial representation of serial order information is further emphasized by the difficulties of deaf individuals in recalling items in a given temporal sequence (reviewed in Rinaldi et al., 2018), but more research is needed to investigate whether this also affects the construction of a number line.

The Case of Space

Space is strikingly different. Not only does it have more dimensions than the other domains under scrutiny, but it also lacks inherent structure, order, and orientation. Apart from a single somewhat privileged direction, defined by gravitation, all other attempts of ordering presuppose a human perspective. Near versus far, front versus back, left versus right all depend on a subjective point of view, and even the non-relativistic reference points that define an absolute frame of reference such as cardinal directions, the slope of mountain sides, or a river's direction of flow require cultural conventions (Levinson, 2003; Bender and Beller, 2014). This may explain why, despite substantial

consistency across other domains, response patterns in the spatial domain do not necessarily converge with any of the others.

Evolution of Alignment Patterns

If this account is valid, the mechanism that may have given rise to spatialized number lines would be less likely a result of a predisposition for a certain type of representations, and more likely a result of cultural evolution (Winter et al., 2015; Núñez, 2017), in the course of which a diverse set of cultural representations emerged that helped us put order into important domains. One of these representations was powerful enough to enable the alignment of ordered sequences across domains. Still, coming up with linear representations is far less trivial than we tend to believe. With the exception of rulers (that actually are an attempt to organize space by way of numbers, rather than the reverse), none of the cultural representations of number or time (and text) is strictly linear, or even linear at all. Number representations, for instance, at least in the decimal systems of the languages under scrutiny here, are 2-dimensional (Zhang and Norman, 1995), and even in the lower range in which they are still 1-dimensional, they are not necessarily represented in a line on a substantial number of cultural devices (e.g., telephone keypads and door locks). Likewise, the Japanese *kana* syllabaries are tabulated to begin with (Figures 2A,B), while the letters of the Latin alphabet in the medium with which we are arguably most frequently confronted are presented neither in a linear nor even an ordered manner (to verify, simply look at your computer's keyboard)—not to mention the innumerable combinations into which they are turned in daily life. Layout for most texts is also 2-dimensional, with lines running primarily left-to-right, but also top-to-bottom on a page. And even time is typically represented cyclically, emphasizing the recurrence of seconds, minutes, and hours (on analog clocks), or in a tabulated manner for both clock-time (on digital clocks) and larger units such as weekdays, months, or years (on calendars). None of these representations should actually prepare people to develop line-like representations, and both historical sources and data on synesthesia attest to some of the ensuing variability (e.g., Galton, 1883; Ernest, 1986; Bender and Beller, 2011; Núñez, 2011).

Arguably the only strictly linear—yet also entirely non-spatial—mode of representation in all domains discussed here (and across languages) is the verbal routine of reciting, be it for the alphabet or the sequence of counting words (the latter initially reinforced by finger counting; cf. Beller and Bender, 2011; Fischer and Brugger, 2011). Once in place, this linear sequence can be harnessed for organizing similarly structured domains such as loudness (Núñez et al., 2011) or time. Space, by contrast, with its three dimensions and lack of inherent structure, defies simple ways of sequencing. While some spatial arrangements do receive order with the help of one of the above domains, such as when hotel rooms or train cars are numbered, most arrangements in the spatial domain necessitate a rather complex coding based on a coordinate system for which both anchoring and aligning need further specification (Levinson, 2003; Bender and Beller, 2014). Taken together, this raises the question of whether it is really (the allegedly more concrete) space that serves as the universal foundation for representations of more

abstract domains such as number and time, or whether inherent features of the latter two—as ordered sequences—are actually what facilitate our organization of space (for examples and other arguments why spatial representations of number or time may not be universal, see also Hutchins, 1983; Núñez, 2011, 2017; Núñez and Cornejo, 2012; Bender and Beller, 2014).

Directions for Future Research

As discussed earlier, standard paradigms for number line assessment obtain spatial representations along the lateral dimension, because they predefine this dimension as part of the task. One of the most important achievements of the current study is therefore its use of a language elicitation task, which allows us to tap into a different (i.e., the sagittal) dimension. Admittedly, the set of tasks in our study also predefines a dimension, even if a different one, in either providing forced-choice response options (e.g., “Number 25 is two positions *in front of/behind* number 23”) or by phrasing the task itself using the dimension under scrutiny (e.g., “Which letter is directly *in front of/behind* G in the alphabet?”). Given our interest in establishing whether the sagittal axis can be used to represent numbers and the fact that such phrasings are more natural in language than left/right or top/down phrasings, we considered this approach justified for an exploratory study. However, if aiming for a more comprehensive understanding of how number representations may principally be aligned with space, future research would be well advised to open up the scope for possible responses. This should include an investigation of whether distinct ways of anchoring, if occurring at all, affect how the number line is spatialized.

A second way in which the current work should be extended lies in the range of languages investigated and the linguistic and cultural factors thus targeted. Specifically, while we attempted to include languages with different writing and reading directions, our selection does not cover the full range of variability in this regard. The same is true for finger counting patterns or properties of counting systems. With regard to the latter, for instance, users of body-based counting systems like the Oksapmin (Saxe, 1981) would be an informative sample. Including more characteristics of cultural groups and language communities may also help to answer the puzzling question of why speakers of closely related languages like the Germanic languages tend to differ so substantially in their mental representations of number and time lines. To this effect, other sample characteristics like differences between dialects or effects of bilingualism would be worth investigating.

In addition, research on people with sensory deprivation would be able to shed more light on which aspects of number representation are possibly innate, which are based on sensorimotor experience of movement, and which are brought about by cultural practices and linguistic routines. Unfortunately, while this line of research is experiencing an upsurge, studies devoted to numerical representations along the sagittal axis are still missing (Rinaldi et al., 2018).

A third possible direction for future research could be a more in-depth investigation of the potential role played by perspective and other task specifics. Apparently, the involvement of motion,

for instance, has the potential to induce a perception of “forward” in the direction of larger numbers, in line with linguistic expressions like “counting forward.” Surprisingly, however, this perspective was observed only in two of the five groups, and even there it was not strong enough to fully reverse the preference for assigning FRONT to the beginning of the number line in static relations. Exactly which factors contribute to the partial reversal of FRONT assignment in some groups, but not in others, therefore remains an open question.

And finally, more research should be devoted to the analysis of changes over time. While we already discussed the impact that cultural and linguistic tools may have had on the emergence and evolution of number line representations, their influence on children’s development deserve similar attention. Apparently, children’s increasing understanding of numbers involves an increasing number of symbolic, culture-specific representations; as a result, the application of procedural knowledge is gradually replaced by the retrieval of declarative knowledge (Sasanguie and Vos, 2018). This raises the interesting question of whether and how increasing knowledge of other cultural systems (e.g., temporal representations or the alphabet) may affect how children learn to represent and process information from those domains, or whether and when generalizations across domains may emerge.

CONCLUSION

Number lines and time lines are an appealing possibility compared to the many other ways in which numbers or dates may be mapped onto spatial representations. However, the high degree of variability in the dimensions or axes recruited and in the orientation of alignment with these axes suggests that no specific linear representation is exclusive or essential. Unless we open up our horizon for alternative possibilities, and amend our toolkit with alternative techniques and tasks, we will not be able to find out which possibilities for representing number, time, and other domains, beyond these simple lines, humans actually possess. People are highly flexible in their representations—and prepared to demonstrate this if only they are provided with respective opportunities. Future research should therefore take this more seriously, both with regard to their theoretical conceptualization and to the designs of research paradigms and tasks.

ETHICS STATEMENT

Although our university ethics board only deals with medical research, we can confirm that we follow the Frankfurt declaration of ethical conduct for anthropological research, which addresses all stages of the research project from designing to reporting the research.

AUTHOR CONTRIBUTIONS

All authors listed have made substantial, direct and intellectual contribution to the work and approved it for publication.

ACKNOWLEDGMENTS

This work was supported partly by the Deutsche Forschungsgemeinschaft (DFG) through a grant for the project *Spatial referencing across languages: Cultural preferences and cognitive implications* to AB (Be 2451/13-1) and SB (Be 2178/7-1) and partly by the Research Council of Norway through the SFF *Centre for Early Sapiens Behaviour (SapienCE)*, project number 262618. For help with data collection and translations, we thank Kristin Sjøfjell and Marleen Wilms (Norwegian data), Megumi Senda, Miriam Seel, Akiko Matsuo, Megumi

Higuchi, and Ami Sato (Japanese data), Lingyan Qian and Wenting Sun (Chinese data), and Elisabeth Kraus (English data). For inspiring discussion and/or valuable comments on earlier versions of this paper, we thank Sarah Mannion de Hernandez.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.01724/full#supplementary-material>

REFERENCES

- Beller, S., and Bender, A. (2011). Explicating numerical information: when and how fingers support (or hinder) number comprehension and handling. *Front. Psychol.* 2:214. doi: 10.3389/fpsyg.2011.00214
- Beller, S., and Bender, A. (2017). "How relative is the relative frame of reference? Front and back in Norwegian, Farsi, German, and Japanese," in *Proceedings of the 39th Annual Conference of the Cognitive Science Society*, eds G. Gunzelmann, A. Howes, T. Tenbrink, and E. J. Davelaar (Austin, TX: Cognitive Science Society), 118–123.
- Beller, S., Bender, A., Chrisomalis, S., Jordan, F. M., Overmann, K. A., Saxe, G. B., et al. (2018). The cultural challenge in mathematical cognition. *J. Numer. Cogn.* 4, 448–463. doi: 10.5964/jnc.v4i2.137
- Beller, S., Singmann, H., Hühner, L., and Bender, A. (2015). Turn around to have a look? Spatial referencing in dorsal vs. frontal settings in cross-linguistic comparison. *Front. Psychol.* 6:1283. doi: 10.3389/fpsyg.2015.01283
- Bender, A., and Beller, S. (2011). Cultural variation in numeration systems and their mapping onto the mental number line. *J. Cross Cult. Psychol.* 42, 579–597. doi: 10.1177/0022022111406631
- Bender, A., and Beller, S. (2012). Nature and culture of finger counting: diversity and representational effects of an embodied cognitive tool. *Cognition* 124, 156–182. doi: 10.1016/j.cognition.2012.05.005
- Bender, A., and Beller, S. (2014). Mapping spatial frames of reference onto time: a review of theoretical accounts and empirical findings. *Cognition* 132, 342–382. doi: 10.1016/j.cognition.2014.03.016
- Bender, A., and Beller, S. (2018). "Numeration systems as cultural tools for numerical cognition," in *Mathematical Cognition and Learning: Language and Culture in Mathematical Cognition*, eds D. B. Berch, D. C. Geary, and K. Mann Koepke (Cambridge, MA: Academic Press), 297–320. doi: 10.1016/B978-0-12-812574-8.00013-4
- Bender, A., Beller, S., and Bennardo, G. (2010). Temporal frames of reference: conceptual analysis and empirical evidence from German, English, Mandarin Chinese, and Tongan. *J. Cogn. Cult.* 10, 283–307. doi: 10.1163/156853710X531195
- Bender, A., Rothe-Wulf, A., Hühner, L., and Beller, S. (2012). Moving forward in space and time: how strong is the conceptual link between spatial and temporal frames of reference? *Front. Psychol.* 3:486. doi: 10.3389/fpsyg.2012.00486
- Bender, A., Sjøfjell, K., Rothe-Wulf, A., and Beller, S. (2017). "Representing time in terms of space: directions of mental timelines in Norwegian," in *Proceedings of the 39th Annual Conference of the Cognitive Science Society*, eds G. Gunzelmann, A. Howes, T. Tenbrink, and E. J. Davelaar (Austin, TX: Cognitive Science Society), 1617–1622.
- Bennardo, G. (2009). *Language, Space and Social Relationships: A Foundational Cultural Model in Polynesia*. Cambridge: Cambridge University Press.
- Bergen, B. K., and Chan Lau, T. T. (2012). Writing direction affects how people map space onto time. *Front. Psychol.* 3:109. doi: 10.3389/fpsyg.2012.00109
- Bonato, M., Zorzi, M., and Umiltà, C. (2012). When time is space: evidence for a mental time line. *Neurosci. Biobehav. Rev.* 36, 2257–2273. doi: 10.1016/j.neubiorev.2012.08.007
- Boroditsky, L. (2000). Metaphoric structuring: understanding time through spatial metaphors. *Cognition* 75, 1–27. doi: 10.1016/S0010-0277(99)00073-6
- Boroditsky, L., and Gaby, A. (2010). Remembrances of times east: absolute spatial representations of time in an Australian aboriginal community. *Psychol. Sci.* 21, 1635–1639. doi: 10.1177/0956797610386621
- Calude, A. S., and Verkerk, A. (2016). The typology and diachrony of higher numerals in Indo-European: a phylogenetic comparative study. *J. Lang. Evol.* 1, 91–108. doi: 10.1093/jole/lzw003
- Casasanto, D., and Jasmin, K. (2012). The hands of time: temporal gestures in English speakers. *Cogn. Linguist.* 23, 643–674. doi: 10.1515/cog-2012-0020
- Dehaene, S., Bossini, S., and Giraux, P. (1993). The mental representation of parity and number magnitude. *J. Exp. Psychol. Gen.* 122, 371–396. doi: 10.1037/0096-3445.122.3.371
- Dehaene, S., Izard, V., Spelke, E., and Pica, P. (2008). Log or linear? distinct intuitions of the number scale in Western and Amazonian indigene cultures. *Science* 320, 1217–1220. doi: 10.1126/science.1156540
- Ernest, P. (1986). Mental number line images. *Teach. Math. Appl.* 5, 1–2. doi: 10.1093/teamat/5.1.1
- Fischer, M. H., and Brugger, P. (2011). When digits help digits: spatial-numerical associations point to finger counting as prime example of embodied cognition. *Front. Psychol.* 2:260. doi: 10.3389/fpsyg.2011.00260
- Fischer, M. H., and Fias, M. H. (2005). "Spatial representation of numbers," in *Handbook of Mathematical Cognition*, ed J. I. D. Campbell (New York, NY: Psychology Press), 43–54.
- Fischer, M. H., Mills, R. A., and Shaki, S. (2010). How to cook a SNARC: number placement in text rapidly changes spatial-numerical associations. *Brain Cogn.* 72, 333–336. doi: 10.1016/j.bandc.2009.10.010
- Fuhrman, O., McCormick, K., Chen, E., Jiang, H., Shu, D., Mao, S., et al. (2011). How linguistic and cultural forces shape conceptions of time: english and Mandarin time in 3D. *Cogn. Sci.* 35, 1305–1328. doi: 10.1111/j.1551-6709.2011.01193.x
- Galton, A. (2011). Time flies but space does not: limits to the spatialisation of time. *J. Pragmat.* 43, 695–703. doi: 10.1016/j.pragma.2010.07.002
- Galton, F. (1883). *Inquiries into Human Faculty and its Development*. New York, NY: Macmillan.
- Göbel, S. M., Shaki, S., and Fischer, M. H. (2011). The cultural number line: a review of cultural and linguistic influences on the development of number processing. *J. Cross Cult. Psychol.* 42, 543–565. doi: 10.1177/0022022111406251
- Grabowski, J., and Miller, G. A. (2000). Factors affecting the use of dimensional prepositions in German and American English: object orientation, social context, and prepositional pattern. *J. Psycholinguist. Res.* 29, 517–553. doi: 10.1023/A:1005124210205
- Grabowski, J., and Weiß, P. (1996). The prepositional inventory of languages: a factor that affects comprehension of spatial prepositions. *Lang. Sci.* 18, 19–35. doi: 10.1016/0388-0001(96)00005-8
- Haun, D. B., Rapold, C. J., Janzen, G., and Levinson, S. C. (2011). Plasticity of human spatial cognition: spatial language and cognition covary across cultures. *Cognition* 119, 70–80. doi: 10.1016/j.cognition.2010.12.009
- Hung, Y. H., Hung, D. L., Tzeng, O. J., and Wu, D. H. (2008). Flexible spatial mapping of different notations of numbers in Chinese readers. *Cognition* 106, 1441–1450. doi: 10.1016/j.cognition.2007.04.017
- Hutchins, E. (1983). "Understanding Micronesian navigation," in *Mental Models*, eds D. Gentner and A. L. Stevens (Hillsdale, NJ: Lawrence Erlbaum), 191–225.

- Ito, Y., and Hatta, T. (2004). Spatial structure of quantitative representation of numbers: evidence from the SNARC effect. *Mem. Cogn.* 32, 662–673. doi: 10.3758/BF03195857
- Kennedy, J. J. (1992). *Analyzing Qualitative Data*. New York, NY: Praeger.
- Lakoff, G., and Johnson, M. (1980). *Metaphors we Live by*. Chicago, IL: University of Chicago Press.
- Le Guen, O., and Pool Balam, L. I. (2012). No metaphorical timeline in gesture and cognition among Yucatec Mayas. *Front. Psychol.* 3:271. doi: 10.3389/fpsyg.2012.00271
- León-Portilla, M. (1990). *Time and Reality in the Thought of the Maya*. Oklahoma, OK: University of Oklahoma Press
- Levinson, S. C. (2003). *Space in Language and Cognition*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511613609
- Majid, A., Bowerman, M., Kita, S., Haun, D. B. M., and Levinson, S. C. (2004). Can language restructure cognition? the case for space. *Trends Cogn. Sci.* 8, 108–114. doi: 10.1016/j.tics.2004.01.003
- Miles, L. K., Tan, L., Noble, G. D., Lumsden, J., and Macrae, C. N. (2011). Can a mind have two time lines? exploring space–time mapping in Mandarin and English speakers. *Psychon. Bull. Rev.* 18, 598–604. doi: 10.3758/s13423-011-0068-y
- Miura, I. T. (1987). Mathematics achievement as a function of language. *J. Educ. Psychol.* 79, 79–82. doi: 10.1037/0022-0663.79.1.79
- Moeller, K., Pixner, S., Kaufmann, L., and Nuerk, H. C. (2009). Children's early mental number line: logarithmic or decomposed linear? *J. Exp. Child Psychol.* 103, 503–515. doi: 10.1016/j.jecp.2009.02.006
- Núñez, R., Cooperrider, K., and Wassmann, J. (2012b). Number concepts without number lines in an indigenous group of Papua New Guinea. *PLoS ONE* 7:e35662. doi: 10.1371/journal.pone.0035662
- Núñez, R. E. (2008). Reading between the number lines. *Science* 321:1293. doi: 10.1126/science.321.5894.1293
- Núñez, R. E. (2011). No innate number line in the human brain. *J. Cross Cult. Psychol.* 42, 651–668. doi: 10.1177/0022022111406097
- Núñez, R. E. (2017). Is there really an evolved capacity for number? *Trends Cogn. Sci.* 21, 409–424. doi: 10.1016/j.tics.2017.03.005
- Núñez, R., and Cooperrider, K. (2013). The tangle of space and time in human cognition. *Trends Cogn. Sci.* 17, 220–229. doi: 10.1016/j.tics.2013.03.008
- Núñez, R., Cooperrider, K., Doan, D., and Wassmann, J. (2012a). Contours of time: topographic construals of past, present, and future in the Yupno valley of Papua New Guinea. *Cognition* 124, 25–35. doi: 10.1016/j.cognition.2012.03.007
- Núñez, R. E., and Cornejo, C. (2012). Facing the sunrise: cultural worldview underlying intrinsic-based encoding of absolute frames of reference in Aymara. *Cogn. Sci.* 36, 965–991. doi: 10.1111/j.1551-6709.2012.01237.x
- Núñez, R., Doan, D., and Nikoulina, A. (2011). Squeezing, striking, and vocalizing: is number representation fundamentally spatial? *Cognition* 120, 225–235. doi: 10.1016/j.cognition.2011.05.001
- Núñez, R. E., and Sweetser, E. (2006). With the future behind them. Convergent evidence from Aymara language and gesture in the crosslinguistic comparison of spatial construals of time. *Cogn. Sci.* 30, 401–450. doi: 10.1207/s15516709cog0000_62
- Rinaldi, L., Merabet, L. B., Vecchi, T., and Cattaneo, Z. (2018). The spatial representation of number, time, and serial order following sensory deprivation: a systematic review. *Neurosci. Biobehav. Rev.* 90, 371–380. doi: 10.1016/j.neubiorev.2018.04.021
- Rothe-Wulf, A., Beller, S., and Bender, A. (2015). Temporal frames of reference in three Germanic languages: individual consistency, interindividual consensus, and cross-linguistic variability. *Q. J. Exp. Psychol.* 68, 917–939. doi: 10.1080/17470218.2014.970205
- Santens, S., and Gevers, W. (2008). The SNARC effect does not imply a mental number line. *Cognition* 108, 263–270. doi: 10.1016/j.cognition.2008.01.002
- Sasanguie, D., Lyons, I. M., De Smedt, B., and Reynvoet, B. (2017). Unpacking symbolic number comparison and its relation with arithmetic in adults. *Cognition* 165, 26–38. doi: 10.1016/j.cognition.2017.04.007
- Sasanguie, D., and Vos, H. (2018). About why there is a shift from cardinal to ordinal processing in the association with arithmetic between first and second grade. *Dev. Sci.* 21:e12653. doi: 10.1111/desc.12653
- Saxe, G. B. (1981). Body parts as numerals: a developmental analysis of numeration among the Oksapmin in Papua New Guinea. *Child Dev.* 52, 306–316. doi: 10.2307/1129244
- Senft, G. (1997). *Referring to Space*. Edn. Oxford: Clarendon.
- Shaki, S., and Fischer, M. H. (2008). Reading space into numbers—a cross-linguistic comparison of the SNARC effect. *Cognition* 108, 590–599. doi: 10.1016/j.cognition.2008.04.001
- Shaki, S., and Fischer, M. H. (2018). Deconstructing spatial-numerical associations. *Cognition* 175, 109–113. doi: 10.1016/j.cognition.2018.02.022
- Shaki, S., Fischer, M. H., and Petrusic, W. M. (2009). Reading habits for both words and numbers contribute to the SNARC effect. *Psychon. Bull. Rev.* 16, 328–331. doi: 10.3758/PBR.16.2.328
- Siegler, R. S., and Opfer, J. E. (2003). The development of numerical estimation: evidence for multiple representations of numerical quantity. *Psychol. Sci.* 14, 237–250. doi: 10.1111/1467-9280.02438
- Sinha, C., Da Silva Sinha, V., Zinken, J., and Sampaio, W. (2011). When time is not space: the social and linguistic construction of time intervals and temporal event relations in an Amazonian culture. *Lang. Cogn.* 3, 137–169. doi: 10.1515/langcog.2011.006
- Torralbo, A., Santiago, J., and Lupiáñez, J. (2006). Flexible conceptual projection of time onto spatial frames of reference. *Cogn. Sci.* 30, 745–757. doi: 10.1207/s15516709cog0000_67
- Tversky, B., Kugelmass, S., and Winter, A. (1991). Cross-cultural and developmental trends in graphic productions. *Cogn. Psychol.* 23, 515–557. doi: 10.1016/0010-0285(91)90005-9
- Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends Cogn. Sci.* 7, 483–488. doi: 10.1016/j.tics.2003.09.002
- Winter, B., Matlock, T., Shaki, S., and Fischer, M. H. (2015). Mental number space in three dimensions. *Neurosci. Biobehav. Rev.* 57, 209–219. doi: 10.1016/j.neubiorev.2015.09.005
- Wood, G., Willmes, K., Nuerk, H. C., and Fischer, M. H. (2008). On the cognitive link between space and number: a meta-analysis of the SNARC effect. *Psychol. Sci.* 50, 489–525. Available online at: <http://psycnet.apa.org/record/2009-00781-003>
- Zebian, S. (2005). Linkages between number concepts, spatial thinking, and directionality of writing: the SNARC effect and the reverse SNARC effect in English and Arabic monoliterates, biliterates, and illiterate Arabic speakers. *J. Cogn. Cult.* 5, 165–190. doi: 10.1163/1568537054068660
- Zhang, J., and Norman, D. A. (1995). A representational analysis of numeration systems. *Cognition* 57, 271–295. doi: 10.1016/0010-0277(95)00674-3

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Bender, Rothe-Wulf and Beller. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Development of a Possible General Magnitude System for Number and Space

Karin Kucian^{1,2,3*}, Ursina McCaskey^{1,2}, Michael von Aster^{1,2,3,4} and Ruth O'Gorman Tuura^{1,2,5}

¹ Center for MR-Research, University Children's Hospital Zurich, Zurich, Switzerland, ² Children's Research Center, University Children's Hospital Zurich, Zurich, Switzerland, ³ Neuroscience Center Zurich, University of Zurich and ETH Zurich, Zurich, Switzerland, ⁴ Clinic for Child and Adolescent Psychiatry, German Red Cross Hospital, Berlin, Germany, ⁵ Zurich Center for Integrative Human Physiology, University of Zurich, Zurich, Switzerland

There is strong evidence for a link between numerical and spatial processing. However, whether this association is based on a common general magnitude system is far from conclusive and the impact of development is not yet known. Hence, the present study aimed to investigate the association between discrete non-symbolic number processing (comparison of dot arrays) and continuous spatial processing (comparison of angle sizes) in children between the third and sixth grade ($N = 367$). Present findings suggest that the processing of comparisons of number of dots or angle are related to each other, but with angle processing developing earlier and being more easily comparable than discrete number representations for children of this age range. Accordingly, results favor the existence of a more complex underlying magnitude system consisting of dissociated but closely interacting representations for continuous and discrete magnitudes.

Keywords: number, space perception, ATOM, magnitude processing, development, angles, children

OPEN ACCESS

Edited by:

Maciej Haman,
University of Warsaw, Poland

Reviewed by:

Robert Reeve,
The University of Melbourne, Australia
Stella Felix Lourenco,
Emory University, United States

*Correspondence:

Karin Kucian
karin.kucian@kispi.uzh.ch

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 07 March 2018

Accepted: 26 October 2018

Published: 19 November 2018

Citation:

Kucian K, McCaskey U,
von Aster M and O'Gorman Tuura R
(2018) Development of a Possible
General Magnitude System
for Number and Space.
Front. Psychol. 9:2221.
doi: 10.3389/fpsyg.2018.02221

INTRODUCTION

Differentiation Between Different Aspects of Number and Space Processing

A strong association between numbers and space has been reported over the last years of research. However, reported findings refer to different aspects of numbers and space. Therefore, it is very important to differentiate between various characteristics of numerical and spatial processing and their interrelation to gain further understanding and disentangle the complex number-space association. In this vein, Patro et al. (2014) proposed a more differentiated discussion of the number-space interaction since different numerical and spatial tasks target different underlying representations. According to their four level system of spatial-numerical associations, the authors suggest two categories with a non-directional number-space mapping: (1) cross-dimensional magnitude processing (number: cardinal, space: non-directional), and (2) association between spatial and numerical intervals (number: interval, space: non-directional). The other two categories refer to directional number-space mapping requiring spatial directionality in a sense that larger numbers are generally associated with the right side in Western cultures, while smaller numbers are associated with the left: (3) associations between cardinalities

Abbreviations: M, mean; N, number; p, statistical p -value; r, Pearson correlation coefficient; SD, Standard Deviation; t, Student's t -test value.

and spatial directions (number: cardinal, space: directional), and (4) associations between ordinalities and spatial direction (number: ordinal, space: directional). The present study focuses on cross-dimensional magnitude processing. This includes the examination of interrelations between cardinal aspects of non-symbolic numerosities (e.g., arrays of dots) and non-directional spatial dimension (e.g., line lengths, angles, sizes). Accordingly, when we talk about or discuss the number-space link in the present study, we exclusively refer to the above-mentioned numerical and spatial characteristics. In detail, number processing was explored by non-symbolic number comparison including two sets of dot clouds and spatial processing by comparison of two angles. Both tasks demand a magnitude judgment, which is based either on the evaluation of discrete quantity estimation of numerosity (number) or on continuous spatial processing (space).

A General Cognitive Magnitude System

Associations between such different kinds of magnitude processing have led to the hypothesis of the existence of a shared general cognitive magnitude representation. Walsh (2003) proposed in “A Theory of Magnitude (ATOM)” that quantity, space, and time are part of a general magnitude system. Recent research has investigated to what extent and why these representational systems are shared. According to the content of the present study, we are mainly providing examples of cardinal numerical and non-directional spatial interactions. For an overview about associations between all dimensions (number, space, time, size, speed) according to ATOM see the review by Bueti and Walsh (2009).

Crucial contributions to the origin and existence of cross-dimensional magnitude processing stem from recent research in infants, brain imaging studies in adults, and single-cell recordings in primates or animals. Different studies highlight that a predisposition to relate numerical information to spatial magnitudes emerges very early in life (de Hevia and Spelke, 2010; Lourenco and Longo, 2010; de Hevia et al., 2012a,b). For instance, de Hevia and Spelke (2010) could show that infants as young as 8 months are sensitive to the association between non-symbolic numerical magnitudes and spatial line lengths. Moreover, also when continuous spatial variables are held constant, infants still attend to numerical change, indicating that number is spontaneously represented by young infants and both spatial and number information are probably integrated in an early magnitude representation (Brannon et al., 2006; Cordes and Brannon, 2009; Starr and Brannon, 2015). Finally, de Hevia et al. (2014) provided evidence that representations of space, time, and number are interrelated in even 0 to 3-day-old neonates.

Studies in adults corroborate a strong relation between number and space on both the behavioral and neuronal levels. Repeatedly a behavioral interference between the judgment of different magnitudes has been reported (Hurewitz et al., 2006; Longo and Lourenco, 2010; Dormal and Pesenti, 2013). On the neuronal level, several studies depicted an overlay of brain activation localized in the parietal lobes for different magnitudes (e.g., Fias et al., 2003; Dormal et al., 2012; for review see

Pinel et al., 2004; Hubbard et al., 2005; Kaufmann et al., 2008). And particularly, the right intraparietal sulcus moved into focus as locus of a possible general magnitude system (for review see Sokolowski and Ansari, 2016). More recently, McCaskey et al. (2017) identified in adolescents the occipito-parietal stream as a common magnitude system for numerical and spatial magnitude comparisons assessed with the same task used in the present study.

Finally, animal behavior suggests that many animal species show a representation of space, number, and time (for review see Gallistel, 1989) and single-cell recordings in primates revealed identical neurons within the posterior parietal cortex that code for discrete non-symbolic numerosities (arrays of dots) and continuous spatial quantity (length) (Tudusciuc and Nieder, 2007).

Taken together, various sources of evidence suggest that number and space are processed by a general magnitude system that is claimed to develop very early in life and comprises identical brain areas of the parietal lobules. However, Bueti and Walsh (2009) emphasize in their latest review that although the parietal lobe may be considered as the “primary magnitude cortex,” it is only one locus of magnitude processing and that there is a magnitude system and not a single magnitude area. Therefore, it is also not surprising that only some activation sites for number, space, and time overlap and a few do not. Furthermore, Bueti and Walsh (2009) point out that an over simplistic view of a general magnitude system would assume systematic interferences between number, space, time and all kinds of magnitudes. This is clearly not the case. In fact, Dormal and Pesenti (2007) reported only an interference effect of space with numbers, whereas Nys and Content (2012) showed the reciprocal interference. Moreover, Hurewitz et al. (2006) demonstrated interference between discrete and continuous stimulus dimensions in both directions. Not only are reported findings inconsistent about the directions of interferences between different magnitudes, Agrillo et al. (2013) and Barth (2008) found absolutely no correlations among non-symbolic estimations (number/space/time or number/space) contradicting the existence of a general magnitude system. Similarly, behavioral and neuronal findings from Cappelletti et al. (2014) also point to distinct systems for quantifying different magnitudes. Their results showed that the proficiency in numerical and continuous quantity tasks was not correlated in participants with a specific math learning disorder (dyscalculia) (e.g., impaired number but spared time and space processing), and moreover, performance in these tasks was partly dissociated in subjects without math problems, both behaviorally and anatomically. Similar findings from populations with specific impairments in one quantitative domain reported preserved abilities in other magnitude domains (Mussolin et al., 2011; Rousselle et al., 2013; Crollen and Noël, 2015). Lourenco et al. (2012) also reported only partly overlapping representations of numerical and spatial magnitudes by showing that number and spatial performance correlated with higher mathematical competence, but number precision contributed uniquely to advanced arithmetic and spatial precision uniquely to geometry in adult subjects. Similar work in children by Lourenco and Bonny (2017), however, revealed no

differentiation between number and spatial performance – the precision of both tasks contributed exactly to the same math measures (calculation and geometry). On the other hand, there is also evidence speaking for a correlation between number and spatial processing, as expected under a general magnitude system. Lourenco and colleagues revealed positive correlation between the performance of comparisons between non-symbolic number arrays and cumulative area in typically developing children (Lourenco and Bonny, 2017) and adults (Lourenco et al., 2012).

Results from DeWind and Brannon (2012) are contradictory to ATOM, which predicts that improving abilities in one domain (e.g., number) would improve other quantitative domains (e.g., space). In this regard, DeWind and Brannon (2012) administered a simple numerical training, reporting an improvement in numerical skills but not in a spatial task. Due to this lacking transfer effects, training only one domain and hoping for improvements in the untrained domain makes no sense. However, interventions focusing on the improvement of the association between number and space are supposed to be more beneficial for basic geometrical and numerical understanding (reviewed by Cipora et al., 2015; Hawes et al., 2017).

In sum, there is no doubt about a strong connection between number and space, however, if both representation originate from a single general magnitude system is contradictory and further research is needed.

Development

An important determinant in the explanation of different findings could be characteristics of investigated populations such as age. Regarding development, findings suggest that we are born with the ability to relate numerical and spatial factors (de Hevia et al., 2014), which probably get further integrated over development as can be observed by directional biases in spatial or numerical line bisection tasks in younger children (7 years of age) to an adult-like behavior in 13-year-old children (van Vugt et al., 2000; Hausmann et al., 2003; Göksun et al., 2013). Hence, it can be inferred from these findings that school-age might be still a critical period in the development of numerical and spatial skills. However, only very little knowledge is available today at this age-range. To our knowledge, only one study examined the relation between spatial and numerical skills over development in school aged children and concluded rather differing mechanisms underlying physical and numerical space in childhood that might integrate in adulthood (Göksun et al., 2013).

Speaking about development, it has to be kept in mind that not only the mere existence of a general magnitude system is disputable, but also different possible developmental trajectories are currently discussed (for review please see Lourenco and Longo, 2011; Lourenco, 2015). According to the classic approach of learning by Gibson and Gibson (1955), the *differentiation view* suggests strongest cross-dimensional associations earlier in life and an increase in differentiation of representations of magnitude dimensions over development. In line with this differentiation view, Newcombe et al. (2015) come to the conclusion in their review that infants begin with a general magnitude system which differentiates into distinct dimensions over developmental time.

In contrast, the *enrichment view* assumes an increase in strength of different magnitude representations over development.

Aim of the Present Study

The goal of the present study was to examine the relation between discrete non-symbolic number processing (arrays of dots) and continuous spatial magnitude processing (angles) taking the important aspect of development into consideration. Therefore, we investigated typically achieving children spanning different school grade levels. Izard and Spelke (2009) have shown that sensitivity to detect relationships of line length and angles shows steady improvement over childhood, reaching asymptote at about 12 years of age. However, the authors also reported differences in the developmental trajectories of length and angle sensitivity, while the sensitivity to length is mature by the age of 8, sensitivity to angle continues to mature until 10. In addition, and as mentioned above, adult-like behavior has been observed in 13-year-old children in spatial or numerical line bisection tasks. Accordingly, the current work focusses on children between 8 and 13 years, as this age range seems to be an interesting developmental stage to test higher cognitive processing of angle and dot comparison. According to Walsh (2003) an interference between both tasks would support ATOM. Regarding development, we expect improvements in numerical and spatial quantitative skills. As the development of numerical and spatial representation is a complex process, different developmental trajectories are possible. The investigation of these developmental courses could provide further evidence for the existence of a general magnitude system or for separate cognitive representations for discrete and continuous magnitudes. On the one hand, a strong cross-dimensional transfer in earlier grade levels and/or parallel development for number and space abilities would support ATOM (proposed by Walsh, 2003). On the other hand, increasing integration among numerical and spatial magnitudes over development and/or dissociated developmental pathways would rather support the idea that quantitative thinking begins with the ability to discriminate between continuous properties. Over development, children learn the correlation between continuous and discrete features suggesting that discrete and continuous magnitude processing are two separate, but interacting systems underlying a general magnitude system (proposed by Leibovich and Henik, 2013a).

To address these hypotheses we decided to test non-symbolic number processing by the comparison of number of dots and spatial magnitude processing by a clearly different stimulus type, namely angle size. This is in contrast to some studies that use exactly the same arrays of dots for both dimensions by asking two different questions: which of two arrays is greater in number (numerical estimation) or cumulative area (spatial estimation). Although such a design has the advantage of using exactly the same stimuli for the two tasks, it has the disadvantage that participants have always to keep in mind which question they have to answer at the moment and even more importantly, they have to inhibit the processing of the irrelevant dimension. Both additional mental processes are not of interest in our study and put a supplementary challenge especially for children. Finally, research with infants proved that they are already able

to discriminate 2-dimensional angles (Slater et al., 1991) and findings from preschool children corroborated generally high performance levels of angle comparisons and provide evidence that the dimension of angle is even more salient than length for children (Izard and Spelke, 2009). Therefore, the present study design testing children's magnitude processing skills uses dot array versus angles.

MATERIALS AND METHODS

Subjects

In total 369 children participated in the present study, of which 2 were excluded due to incomplete task performance, resulting in a group size of 367 children between 8.2 and 12.9 years of age ($M = 10.6$; $SD = 1.1$), including 39% girls and 61% boys. Children attended third to sixth school grades, such that 87 children were in the third grade (8.2–10.2 years of age: $M = 9.3$; $SD = 0.4$), 140 in the fourth grade (9.3–11.8 years of age: $M = 10.3$; $SD = 0.4$), 110 in the fifth grade (10.1–12.7 years of age: $M = 11.4$; $SD = 0.5$), and 30 in the sixth grade (11.0–12.9 years of age: $M = 12.3$; $SD = 0.4$).

The study was approved by the local ethics committee (Kantonale Ethikkommission Zürich) based on guidelines from the World Medical Association's Declaration of Helsinki (WMA, 2002). According to the local ethical committee, written parental consent was not required as no risk for the children existed, voluntariness and privacy was guaranteed at all times. Data collection was fully anonymised and took place in the scope of a lecture of the Children's University of Zurich to illustrate our research field, research question, and research experiments. Children's University of Zurich gave also their consent to analyze obtained data.

Non-symbolic Number Comparison Task

Non-symbolic number comparison performance was tested with a paper-and-pencil task including a total number of 28 different trials (see **Figure 1A**). In each trial two groups of dots including a range from a minimum of 8 to a maximum of 32 dots were presented horizontally. Children were asked to indicate on which side more black dots were presented. Presentation of dots was controlled for individual size of dots (no judgment possible due to individual dot size), total displayed area (no judgment possible due to total black area), distribution of dots (no judgment possible due to total covered area), the total number of presented dots for each numerical distance between sets (control for size effect), the side of correct answer, and comparable number of trials for each numerical distance between presented magnitudes were presented (distance 2 = 4 trials, distance 4 = 4 trials, distance 6 = 6 trials, distance 8 = 5 trials, distance 10 = 5 trials, distance 12 = 4 trials). Ratio between smaller and larger dot arrays was 0.4, 0.5, 0.6, 0.63, 0.67, 0.70, 0.71, 0.77, 0.8, 0.83, 0.9, or 0.91. Detailed information about all 28 trials can be found in the **Supplementary Table S1**. All children were carefully introduced to the task and encouraged to solve all trials by comparison of both sets of presented dots by numerical estimation and not counting. To further prevent children from counting, time was restricted to 2 min for all 28 trials. The ability of non-symbolic

magnitude comparison by dots requires a decision about discrete quantity.

Spatial Comparison Task

In the spatial comparison task, a green and a blue Pacman facing to the right side with varying mouth size was presented horizontally (see **Figure 1B**). Children had to indicate by pencil which of the two presented Pacmen has a bigger mouth, whereas line length intersecting the angle was controlled and corresponded always to radius of the circle. In contrast to the non-symbolic number comparison task, this task requires a visuo-spatial and continuous magnitude decision. The mouth angle of one Pacman was always 45 degrees and the mouth angle of the other Pacman varied between minimum 18 degrees to maximum 72 degrees [18, 23, 27, 32, 36, 40, 42, 47, 49, 54, 59, 63, 68, 72 degrees (2 trials for each degree)]. Difficulty level was controlled by varying the ratio between both presented mouth angles across all trials. Detailed information about all 28 trials can be found in the **Supplementary Table S2**. In addition, the side of the correct answer and color of Pacman were balanced. Similar to the number comparison task, children were carefully instructed and advised to solve the spatial comparison task by simple estimation of mouth sizes and not to use for instance their fingers or any other tool to measure the mouth sizes. Again, children had 2 min time to solve all 28 trials.

Data Analyses

For both tasks, the non-symbolic number and the spatial comparison task, the percentage of correctly solved trials was calculated. Subsequent statistical analyses were performed with IBM SPSS Statistics Version 22. As accuracy levels of number and spatial comparisons were negatively skewed and the assumption of normality was therefore violated, non-parametric tests were used. First, we were interested to see which task is more difficult. Therefore, the percentage of correctly solved trials between both tasks was compared by the Wilcoxon signed-rank test. Subsequently, *post hoc* Wilcoxon signed-rank comparisons between both tasks were performed for each grade level individually. Second, to test if numerical and spatial processing are related, Spearman's correlation coefficients were calculated between both tasks over all grade levels and for each grade level individually. Third, development across grade levels was evaluated by Kruskal-Wallis test, and the *post hoc* Mann-Whitney test was conducted to test for developmental differences between grade levels. Finally, effect sizes are reported for all major findings with the denomination r for dependent Wilcoxon tests and Spearman's correlations, and the denomination q that permits to interpret the difference between two correlations.

RESULTS

All 367 children were able to solve all 28 trials of both tasks within the allotted time of 2 min for each condition and performed clearly above chance level. The median accuracy for the non-symbolic number task ranged from 61–100% across grade levels (third grade $Mdn = 92.9$ (IQR 89.3–96.4); fourth

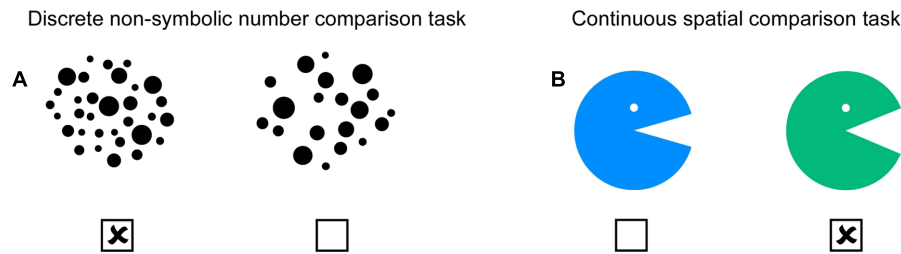


FIGURE 1 | Numerical and spatial task. **(A)** In the non-symbolic number comparison task, children were asked to indicate which set includes more dots. **(B)** In the spatial comparison task, children had to mark the Pacman with the bigger mouth. Both tasks consisted of 28 different pairs, each of which were controlled for factors not of interest. Children had 2 min time for each task to tick the correct solutions. All trials of a task were printed on a double-sided A4 page.

grade Mdn = 96.4 (IQR 92.9–96.4); fifth grade Mdn = 96.4 (IQR 92.9–100); sixth grade Mdn = 96.4 (IQR 92.9–100). Similarly, the median accuracy for the spatial comparison task ranged from 57.1–100% (third grade Mdn = 92.9 (IQR 89.3–96.4); fourth grade Mdn = 92.9 (IQR 89.3–96.4); fifth grade Mdn = 92.9 (IQR 89.3–96.4); sixth grade Mdn = 96.4 (IQR 92.9–100). Please see **Figure 4**.

For any statistical comparisons between both magnitude dimensions, only identical ratios were included in the analyses to prevent any confounding effects due to subtle differences in ratios between tasks. Examining only trials with matched ratios in both conditions resulted in 25 different trials for the number task and 24 trials for the space task. Ratios included in this balanced analysis were as follows: 0.4/0.5–0.51/0.6/0.63/0.66–0.67/0.71/0.76–0.077/0.8/0.83/0.89–0.9/0.91–0.92. Please see **Supplementary Table S3** for detailed information. Including

only matched ratios, the median accuracy for the non-symbolic number task ranged from 60–100% across grade levels [third grade Mdn = 92 (IQR 88–96); fourth grade Mdn = 96 (IQR 92–96); fifth grade Mdn = 96 (IQR 92–100); sixth grade Mdn = 96 (IQR 92–100)]. Similarly, the median accuracy for the spatial comparison task ranged from 63–100% [third grade Mdn = 95.8 (IQR 91.7–100); fourth grade Mdn = 100 (IQR 95.8–100); fifth grade Mdn = 97.9 (IQR 95.8–100); sixth grade Mdn = 100 (IQR 95.8–100)]. Please see **Figure 2**.

Number or Space Comparison: Which Task Is More Difficult?

Results of the Wilcoxon test for identical ratios revealed that spatial comparison (Mdn = 100) is generally easier ($z = -6.771$, $p < 0.001$, $r = -0.25$, $N = 366$) compared to non-symbolic

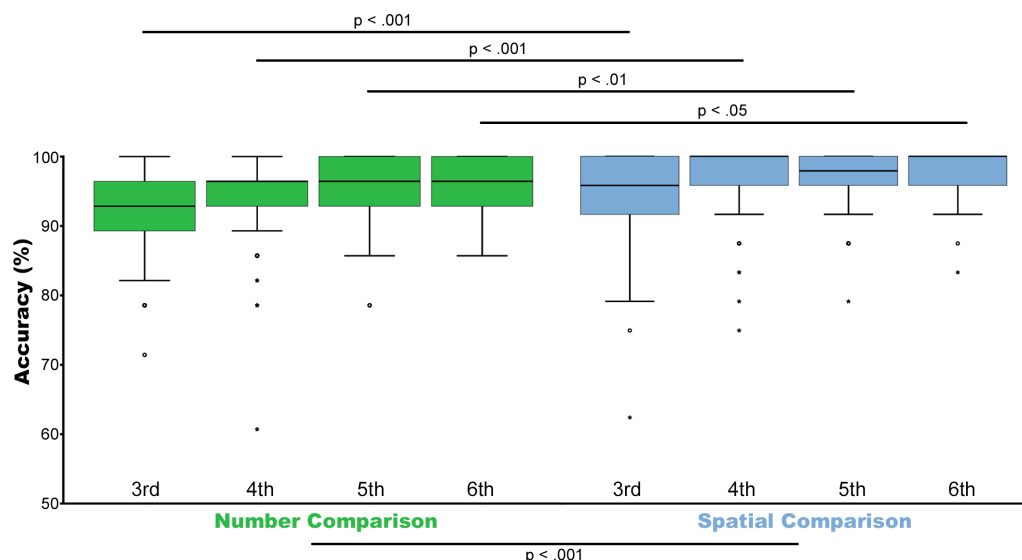


FIGURE 2 | Accuracy. Illustrated are median, interquartile range (IQR = length of box) and lowest and highest values which are no greater than 1.5 times the IQR (whiskers) of percentage correctly solved trials for non-symbolic number comparison (green) and spatial comparison (blue) from the third to the sixth grade. Outliers are marked by circles (1.5–3 times the IQR from the quartile) or asterisks (a value > 3 times the IQR from the quartile). Wilcoxon test showed that spatial comparison is in general significantly easier compared to non-symbolic number comparison ($p < 0.001$). Analyses between individual grades indicated difference between the number and spatial task in the third ($p < 0.001$), fourth ($p < 0.001$), fifth grade ($p < 0.01$), and sixth ($p < 0.05$) grade. Only trials with matched ratios between conditions were included.

number comparison ($Mdn = 96$). Analyses between tasks for different grade levels individually revealed significant difference between the number and spatial task in the third $z = -3.534$, $p < 0.001$, $r = -0.27$, $N = 87$; fourth grade $z = -4.940$, $p < 0.001$, $r = -0.29$, $N = 139$; fifth grade $z = -2.7$, $p < 0.01$, $r = -0.18$, $N = 110$; and sixth grade $z = -2.083$, $p < 0.05$, $r = -0.27$, $N = 30$. Please see **Figure 2**.

In addition **Figure 3** illustrates that accuracy levels decreased significantly for both conditions with increasing ratio between magnitudes, whereas bigger ratios stand for smaller distances between magnitudes and are therefore more difficult to compare (Spearman's correlation for number comparison: $r_s = -0.961$, $N = 25$, $p < 0.001$; and for spatial comparison: $r_s = -0.880$, $N = 24$, $p < 0.001$). Differences between conditions for different ratios did not reach significance. Please see **Figure 3**.

Are Non-symbolic Number and Spatial Abilities Related?

Spearman's correlation over all grade levels showed that the accuracy for both tasks with matched ratios are significantly and positively related with each other $r_s = 0.264$, $N = 366$, $p < 0.001$, also when partialling out age ($r = 0.257$, $N = 363$, $p < 0.001$) or grade level ($r = 0.250$, $N = 363$, $p < 0.001$) or age and grade together ($r = 0.247$, $N = 362$, $p < 0.001$). *Post hoc* analyses within grade levels supported a relation between both magnitude dimensions. Significant and positive correlations between the number and spatial task were also found within third, fourth, and sixth grade (third grade $r_s = 0.295$, $N = 87$, $p < 0.01$; fourth grade $r_s = 0.305$, $N = 139$, $p < 0.001$; sixth grade $r_s = 0.386$, $N = 30$, $p < 0.05$), but not within fifth (fifth grade $r_s = 0.124$, $N = 110$, $p = 0.196$).

Further, we were interested to evaluate if the strength of correlation between both tasks decreases with development, as the analyses of correlations between both tasks for each grade level pointed into this direction. Therefore, we performed comparison of correlation coefficients between grade levels, using Fisher r -to- z transformation. This revealed significant differences between correlation coefficients of number and space between third and fifth grade (one-tailed $p < 0.05$, Cohen's $q = 0.267$) and between fourth and fifth grade (one-tailed $p < 0.01$, Cohen's $q = 0.337$), pointing to a weaker correlation between magnitude dimensions in fifth grade compared to lower grades. Comparisons between the strength of correlations between number and space of the sixth grade and lower grades turned out not to reach significance.

Development of Non-symbolic Number and Spatial Skills

The developmental course from the third to the sixth grade level of both tasks was evaluated by Kruskal–Wallis test including the performance of all ratios of the two tasks (non-symbolic number task and spatial comparison task) as dependent variable and grade level as independent variable. Results indicated that only for non-symbolic number comparison a significant developmental effect over grade levels could be observed [$H(3) = 15.688$, $p < 0.005$], but not for spatial comparison

performance [$H(3) = 6.848$, $p = 0.77$]. *Post hoc* Mann–Whitney test comparison for non-symbolic number comparison performance showed a significant difference between third and fourth ($U = -1.980$, $p < 0.05$), third and fifth ($U = -3.235$, $p < 0.01$), third and sixth ($U = -3.364$, $p < 0.01$), and between fourth and sixth ($U = -2.079$, $p < 0.05$) grade levels. Please see **Figure 4**.

Comparable developmental effects were found when calculating Spearman's correlations between task performance, age, and grade level. Only non-symbolic number comparison correlated significantly with age ($r_s = 0.157$, $N = 367$, $p < 0.01$) and grade level ($r_s = 0.205$, $N = 367$, $p < 0.001$). Spatial comparison did not reach significance with age ($r_s = 0.034$, $N = 364$, $p = 0.514$), or grade level ($r_s = 0.063$, $N = 364$, $p = 0.229$).

DISCUSSION

The present study aimed to further elucidate the association between number and space, which has been proposed to rely on a common general magnitude system (Walsh, 2003; Buetti and Walsh, 2009). However, conflicting research findings called into question whether processing of different dimensions of magnitudes can be attributable to such a general magnitude system. To extend the current body of literature, we investigated the relationship between discrete non-symbolic number processing and continuous spatial magnitude encoding, taking the impact of development into consideration. To our knowledge, this represents the first attempt to investigate a developmental association between these quantity skills in children between the third and sixth grade. Discrete non-symbolic number processing was tested by means of a comparison task of arrays of dots and continuous spatial processing by the comparison of angle sizes.

In sum, results indicated that angle comparisons were generally easier compared to non-symbolic numerical comparisons for children between the third and sixth grade. Moreover, the larger the ratio between magnitudes that had to be compared the more difficult both conditions became. Second, both tasks were significantly related with each other over the entire examined age range, also when controlling for age and/or grade level. However, third, and lastly, our findings suggest differences in the developmental course of discrete and continuous magnitude processing: significant improvements of discrete numerical processing from the third to the sixth grade can be found, whereas continuous spatial representation might have already reached ceiling levels at this age range.

Number or Space Comparison: Which Task Is More Difficult?

Overall, both tasks got more difficult as the ratio between the two magnitudes increased. This is consistent with the well described distance and size effects which are characterized by increasing difficulty with smaller numerical distances and the larger total numbers of dots to be compared (Moyer and Landauer, 1967). Both effects can be explained by the assumption that our representation of quantitative dimensions become increasingly imprecise and noisy with increasing magnitudes

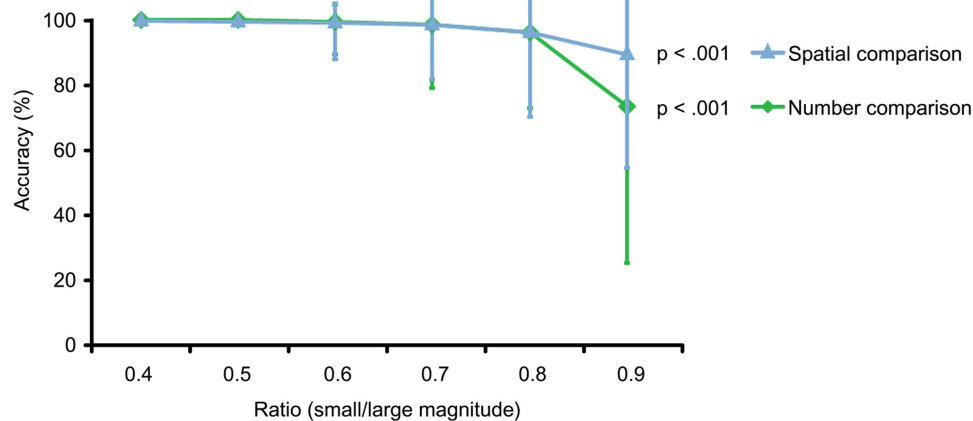


FIGURE 3 | Ratio effect. With increasing ratio between magnitudes, task difficulty increases for both tasks, which is reflected in decreasing accuracy levels for spatial comparison (blue) $p < 0.001$ and number comparison (green) $p < 0.001$. Illustrated are medians and interquartile ranges for each ratio. Only trials with matched ratios between conditions were included.

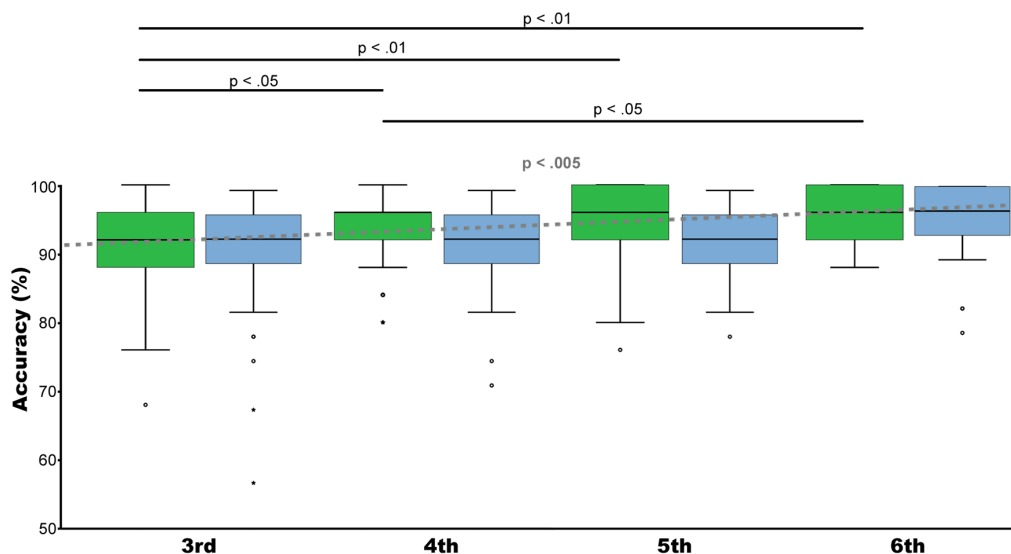


FIGURE 4 | Development. Illustrated are median, interquartile range (IQR = length of box) and lowest and highest values which are no greater than 1.5 times the IQR (whiskers) of percentage correctly solved trials for non-symbolic number comparison (green) and spatial comparison (blue) from the 3rd to the 6th grade. Kruskal–Wallis test showed an increase in mean accuracy over grade levels only for non-symbolic number comparison ($p < 0.005$) (gray dotted line). *Post hoc* analyses revealed significant performance differences between 3rd and 4th ($p < 0.05$), 3rd and 5th ($p < 0.01$). Third and 6th ($p < 0.01$) grade and between 4th and 6th grade ($p < 0.05$). Trials of all ratios were included.

(Feigenson et al., 2004). This seems to be true for both, discrete non-symbolic magnitude and angle processing.

It is important to take into account that ratios between smaller and larger magnitudes differed slightly between both conditions. This might lead to difficulty differences between tasks due to task design in favor to the non-symbolic number comparison task. Therefore, we included only trials with identical ratios of both conditions in order to compare accuracy levels between conditions. Results showed that for matched ratios between tasks spatial judgment of angle size is easier compared to non-symbolic magnitude comparison. This result is consistent with

findings from Leibovich and Henik (2013a), who also showed higher accuracy levels for a continuous spatial task compared to non-symbolic dot comparison. They hypothesize that the superiority of processing continuous magnitudes, together with the fact that evolutionary ancient species such as fish and bees are also able to process continuous magnitudes, might indicate that the system for continuous magnitudes is older than the system for processing discrete magnitudes. Our data can lend support to this assumption that the system for continuous quantity might develop earlier during childhood than the discrete quantity system. In addition, present findings are in line with

results by Odic et al. (2013), who also showed higher acuity for area representation than number representation in 3- to 6-year-old children by comparison of discrete non-symbolic number processing (comparison of dot arrays) and continuous spatial processing (comparison of area sizes).

In addition, both tasks were constructed in a way that confounding factors, such as visual cues, could be excluded to a large degree since many studies have shown that especially in non-symbolic dot comparison tasks results could be biased by visual perceptual cues (e.g., Gebuis and Gevers, 2011; Cleland and Bull, 2015).

Taking these considerations into account, the present findings demonstrate that continuous spatial judgments seems to be easier for school children between third and sixth grade than non-symbolic number discrimination.

Finally, consideration should be given to general angle perception. In the present study, we have assumed that spatial processing of continuous angles is similar to other types of continuous spatial functions, whereas comparison between two angles is getting more difficult the closer both angles are (distance effect) and is getting more difficult with increasing angle sizes for a given distance between angles (size effect). In line with the distance effect, angle comparison got more difficult as the ratio between the two angles increased. However, future studies should test size effects in angle perception particularly.

Are Non-symbolic Number and Spatial Abilities Related?

Present findings reveal that non-symbolic number processing is positively related to continuous spatial abilities in school children. However, since performance of number comparison increased significantly over age or grade levels this relation might have been driven rather by developmental processes. This possibility could be excluded by controlling the effects of age and/or grade level and additionally, correlations between both tasks were also found for third, fourth, and sixth grade level separately. This is in line with behavioral reports of significant interference between numerical and spatial processing in adults (Hurewitz et al., 2006; Longo and Lourenco, 2010; Dormal and Pesenti, 2013). In particular, both Hurewitz et al. (2006) and Dormal and Pesenti (2013) also reported a link between non-symbolic number comparison and continuous spatial processing in adults. Although research in childhood provides strong evidence of mapping numbers and space on a mental number line with a particular scaling (e.g., Siegler and Opfer, 2003; Siegler and Booth, 2004; Booth and Siegler, 2008; Moeller et al., 2009; Kucian et al., 2011) and direction (e.g., Patro and Haman, 2012; Hoffmann et al., 2013; Ebersbach et al., 2014), very little and critically discussed knowledge about children's representation of symbolic numerosities and continuous space is available (de Hevia and Spelke, 2009; de Hevia, 2011; Gebuis and Gevers, 2011; Göksun et al., 2013; Odic et al., 2013; Cleland and Bull, 2015). To our knowledge, no study in terms of discrete non-symbolic numerical quantities in respect to continuous space mapping in school aged children exists to date. Therefore, the present findings could extend the current limited body of literature in school children, showing

an interrelation between cardinal aspects of non-symbolic numerosities and non-directional spatial dimension processing. In contrast, Odic et al. (2013) could not find a significant correlation between number and area acuity in their sample of children once age was controlled. Hence, their results favor rather separate representations of number and area. Contrasting findings might be due to age differences, since children in Odic et al. (2013) study were much younger (3–6 years) compared to the present study (8–13 years). However, it has to be mentioned that reported findings in younger children are mixed and for instance Lourenco and Aulet (2018) reported in a recent study cross-magnitude interactions in infancy and at 3.7 years of age.

Comparison of correlation strength between different grade levels pointed rather to weaker relation between numerical and spatial representations in the fifth grade compared to lower grades. This lower strength of cross-dimensional correlation might hint to an increasing differentiation among magnitude dimensions from third to fifth grade, favoring the differentiation view of development. Similarly, Lourenco et al. (2012) reported a differentiated relation of numerical and spatial magnitude processing on arithmetic and geometry in adulthood, whereas no such differentiation was found in children (Lourenco and Bonny, 2017). However, in the present study, the tendency from third to fifth grade could not be extrapolated into the sixth grade. The comparison of correlation strength of sixth graders with lower grades reached not significance. This might be an effect of the smaller sample size in the sixth grade ($N = 30$). Correlations calculated on data collected from a small sample (30 or fewer subjects) can be affected substantially by dissimilar distribution shapes (Goodwin and Leech, 2006). Whereas in larger sample sizes, there is no direct bearing of sample size to the size of the correlation coefficient (Goodwin and Goodwin, 1999). Accordingly, it might be possible that the examination of a larger sample in the sixth grade would lead to significant differences in correlation strength between sixth and lower grades. However, we cannot tell if correlation strength in sixth grade would be weaker or stronger compared to lower grades. Fact is, that we found no differences between sixth grade and lower grades in the present study and therefore when we consider the total examined age range from third to sixth grade, present data does not legitimate a conclusion in favor of the differentiation or the enrichment view. Future studies should use identical stimuli, as in the present work, in broader age ranges with comparable sample sizes to test the hypothesis of differentiation between magnitude dimensions, because it might also be possible that decreased correlation between dimensions is due to increased general task performance over development.

Development of Non-symbolic Number and Spatial Skills

Various attempts have been made to investigate the development of non-symbolic number processing, but far fewer have examined the development of continuous spatial skills in children. The present study allows not only insights into the developmental course of both skills to be

drawn, but also lends insight into the relationship between them.

Regarding number development, our results are in line with existing knowledge showing that children became more accurate when two non-symbolic magnitudes have to be compared with increasing grade level or age. The nature of children's non-symbolic magnitude representation is thought to index the precision to process numerical quantity information in an approximate way (Dehaene et al., 1999; De Smedt et al., 2013). Halberda and Feigenson (2008) showed that the resolution of this system continues to increase throughout childhood – children perform more accurately and faster on magnitude comparison tasks with increased precise representation.

However, it has also to be taken into account that non-symbolic magnitude comparison tasks in an experimental design as used in the present and other scientific studies control for as many visual nuisance factors as possible to prevent that subjects are able to base their magnitude judgment not on number, but on other magnitude dimensions such as spatial extent. At the same time, such a controlled presentation of non-symbolic magnitudes does usually not correspond to natural surroundings, where more apples take up more space. Therefore, it needs cognitive demands to suppress irrelevant magnitude dimensions in a controlled experimental setting and hence, increased performance in non-symbolic discrete magnitude comparison might also be explained by increased abilities of children in these rather domain-general cognitive capacities than pure numerical abilities (Szucs et al., 2013).

In contrast to numerical magnitude judgment, our findings suggest that from third to sixth grade, children seem not to improve in continuous spatial processing, which is indicated by no correlation between spatial performance levels and age or grade level. Accordingly our data rather indicate no improvements over developmental time in the capacity to compare continuous spatial dimensions at this age range. Alternatively, these data might also be interpreted in a way that discrete numerical magnitude representation is still developing from the third to the sixth grade, whereas continuous spatial processing already reached ceiling level in this age range. In line, Odic et al. (2013) reported similar growth pattern across development for number and area processing in preschool children, but with improvements in area acuity occurring more quickly than in number acuity. The authors argue that these results suggest both an underlying similarity and an important difference between discrete non-symbolic number processing and continuous spatial processing.

General Magnitude System

At large, the present study aimed to gain knowledge about the relation between discrete non-symbolic number encoding and continuous spatial magnitude processing accounting for developmental effects. To date, research has revealed a largely inconclusive picture with respect to an underlying common magnitude system to process both quantity dimensions.

Regarding ATOM, it has been proposed that children with difficulties in one quantitative domain, e.g., numerical processing, should have difficulties in all magnitude domains,

e.g., spatial and temporal encoding. Applied to our study, a child with problems in the number task should also be weak in the spatial task, resulting in equal performance levels between both quantitative tasks. However, in our point of view the mere difference in accuracy levels is a very weak indicator of the relation between two tasks and does not justify the support or contradiction of ATOM. Moreover, it is possible that a single processing system is more prone to one or the other input modality, e.g., due to familiarity, leading to performance differences. In the same vein, acuity of a given magnitude depends on the format of the stimuli, and differences in accuracy levels between different stimuli types, as used in the present study, are probably driven by the stimuli type and not explicable by different magnitude representations (Price et al., 2012; Gilmore et al., 2015). In this sense, interferences such as transfer effects of training one competence on another (as has been carried out for instance by DeWind and Brannon, 2012), priming effects or correlative analyses between tasks are more meaningful.

In the present study, correlation analyses between both tasks pointed to a relation between number and space processing. This link was independent of age or grade level, as the correlation between number and space was still significant when controlling for both factors. Accordingly, we can conclude that discrete non-symbolic number processing and continuous spatial processing are related in school-aged children, but if both skills are processed by a single magnitude system or by two closely interacting systems remains unclear. However, when we take observed differences in the developmental courses of number and space processing into account, the present study provides stronger evidence for two dissociated, but closely related magnitude systems.

On the grounds of current literature and present findings, the description of ATOM as initially proposed by Walsh (2003) seems to be over simplistic as also pointed out by Buetti and Walsh (2009) themselves. Present findings favor suggested models by Leibovich and colleagues (Leibovich and Henik, 2013a,b; Leibovich et al., 2017), who postulate that we are born with the ability to discriminate between continuous properties. As a matter of fact, continuous and discrete properties of arrays of dots for instance are inseparably linked (for review see Leibovich and Henik, 2013b). This is also the case in the present study. Although we have tried to control as many visual confounds as possible in the non-symbolic magnitude comparison task, such as the total surface area of the dots, size of dots, their density, etc., the arrays always contain continuous properties as well. Non-symbolic number comparisons always carry continuous properties that are correlated with numerosities and a separation is physically not possible. Consequently, over development we learn the correlation between continuous and discrete features, which allows us to use both properties to estimate magnitudes. In line with their assumption, our results point to developmental differences of continuous spatial and discrete non-symbolic number processing, with continuous representation being sufficiently developed in third grade children. In contrast, discrete number estimation is still developing and generally more demanding for school children. Moreover, Leibovich and Henik (2013a) suggest that discrete and continuous magnitude

processing are two separate, yet interacting systems underlying a general magnitude system (see also Leibovich et al., 2013). Similarly, current findings showed a link between both number and space processing, also when controlling for age and/or grade level effects, supporting an interaction between systems. On the other hand, differences in general performance levels and developmental trajectories found in the present study also point to partly independent systems. Such a complex interrelated representation of space and number might also explain why the ability to create number–space connections provides only limited links to mathematical learning (reviewed by Cipora et al., 2015). Moreover, Lourenco and Longo (2011) emphasized that characterizing the development of a general magnitude system is complicated and developmental accounts, which consider only differentiation or integration of different magnitudes over time are likely to be incomplete.

Finally, as we have emphasized in the introduction section, it is very important to differentiate between various characteristics of numerical and spatial processing and their interrelation to gain further understanding and disentangle the complex number–space association. In particular, many studies examine the comparison of dot arrays (as in the present study) with area, total cumulative area or line length. In this sense, the present findings add further knowledge on another dimension of continuous spatial processing, namely angles. Accordingly, differences in stimuli type should be considered in the interpretation of different findings. Future research is needed to particularly investigate the relation of discrete non-symbolic number comparison with a variety of continuous types of spatial judgments (area, total cumulative area, angle, length, etc.) to gain a differentiated picture about their relations over development and a possible underlying general magnitude system.

Limitations

As mentioned earlier, present findings are not able to explain the principle of a possible general magnitude system conclusively and some limitations have to be considered. First, although there is lots of evidence showing a relation between different magnitude dimensions, which has been argued to originate from a common general magnitude system, also other explanations for such a crosstalk are possible. Van Opstal and Verguts (2013) for instance propose instead of a general magnitude system that different magnitude representations are processed separately, but share a decision/response procedure or working memory demands which lead to observed similarities between different magnitude dimensions. Similarly, we are not able to distinguish if errors in either task are based on difficulties in number and/or spatial processing or are rather a result of diminished executive functions, like reduced attentional or inhibitory control. However, as our task required no working memory, a relation between dimensions based on common working memory procedures can be excluded. Moreover, the expected and observed increase in difficulty with increasing ratio between sets also speaks against effects of general decision/response procedures or differences in executive processes. Nevertheless, future studies examining numerical and angle processing with a task (e.g., habituation task or priming task) that

is not dependent on domain-general functions and does not require a decision or a response would provide more information regarding this debate. For the relation between non-symbolic numerosity and total cumulative area, Lourenco et al. (2016) tested transfer effects across magnitudes in a subliminal priming paradigm. Their findings suggest that number and area are not fully differentiated, as primed numerals had an effect on performance of cumulative area judgments.

Second, it has to be noted that the present study served as survey of children's non-symbolic numerical and spatial magnitude discrimination abilities to develop a sophisticated paradigm examining also underlying neuronal processes (McCaskey et al., 2017). This is the reason why continuous ratios for numerical and spatial comparisons were chosen to map children's performance levels as thoroughly as possible, but included also slightly different ratios between dimensions. Therefore, it is mandatory to include only identical ratios of number and space judgments as soon as you do any comparison between both magnitude dimensions. Accordingly, we performed separate analyses, including only matched ratios between both tasks to draw clear-cut conclusions regarding comparisons between magnitude dimensions. Correspondingly, **Figure 2** including only matched ratios illustrates higher accuracy levels for spatial comparisons. In contrast, when including all ratios this effect seemed to be reversed, please see **Figure 4**. However, this is falsified by the fact that the spatial comparison task included more trials with higher ratios, which are more difficult to be compared. Therefore, it is important to compare difficulties between conditions only for identical ratios.

Third, performance levels were generally high, why possible ceiling effects have to be considered. However, non-parametric statistical analyses revealed significant differences between both tasks, even when controlling for age or grade level effects, corroborated that difficulty increases with smaller distances between magnitudes that had to be compared, and finally showed improved performances from third to sixth grade for number comparison. None of these effects would be expected if strong ceiling effects were present. However, decreased strength of correlation between number and space from third to fifth grade might be explained by increased general performance up to ceiling levels.

Finally, although children were instructed to compare the angles between both Pacmen, they might solved the task instead by comparing the distance between both mouth sides. In other words, they compared length instead of angles. Since both dimensions depend on continuous spatial processing no differences are expected (please see also Fias et al., 2003). Moreover, many studies use the comparison of length to examine continuous spatial representation (de Hevia and Spelke, 2010; de Hevia et al., 2012b; Dormal and Pesenti, 2013). However, in the present study, children were instructed to compare angles and it can be assumed that the majority did pay attention to angles and not to line length. A further advantage of angle comparison is the similarity to dot comparison as both tasks need spatial processing

in two-dimensions that comprise comparable spatial extent. In contrast, spatial line length elongation is smaller compared to the spatial extent of arrays of dots why angle comparison is favored in the present study.

Finally, to gain a clearer picture of the developmental trajectories of continuous and discrete magnitude processing, future studies should also investigate younger children and measure reaction times to obtain a finer and continuous dimension of performance levels. Moreover, it would be very interesting to relate reaction times or accuracy levels to individual basic numerical and mathematical skills. To do so, future studies should assess a wide range of basic numerical and mathematical abilities that rely more or less on visuo-spatial magnitude processing and relate these skills to individual continuous and discrete magnitude functions.

CONCLUSION

Research has revealed a largely inconclusive picture with respect to the association between numerical and spatial magnitude processing and a common underlying general magnitude system. Our findings provide new insights about the relation of discrete non-symbolic number processing (comparison of dot arrays) and continuous spatial processing (comparison of angle sizes) in children from the third to the sixth grade. Specifically, our results suggest that continuous spatial and discrete number processing are related to each other, but that continuous spatial representations might develop earlier than discrete number representations and are easier to be compared for children at this age range. In conclusion, present findings favor the existence of a more complex underlying magnitude system consisting

of dissociated but closely interacting parts for continuous and discrete magnitude processing.

AUTHOR CONTRIBUTIONS

KK and UM conceived and planned the experiments and tested all the children. KK prepared all the data, performed the statistical data analyses, and wrote the manuscript. UM, MvA, and RO'GT provided the critical feedback to the final version of the manuscript.

FUNDING

This work was supported by the Center for MR-Research of the University Children's Hospital and the NOMIS Foundation, Switzerland.

ACKNOWLEDGMENTS

We would like to thank the Children's University of the University of Zurich for their support in data acquisition and special thanks goes to Marisol Fernandez Cueli for her contribution in data analyses.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.02221/full#supplementary-material>

REFERENCES

- Agrillo, C., Piffer, L., and Adriano, A. (2013). Individual differences in non-symbolic numerical abilities predict mathematical achievements but contradict ATOM. *Behav. Brain Funct.* 9:26. doi: 10.1186/1744-9081-9-26
- Barth, H. C. (2008). Judgments of discrete and continuous quantity: an illusory stroop effect. *Cognition* 109, 251–266. doi: 10.1016/j.cognition.2008.09.002
- Booth, J. L., and Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Dev.* 79, 1016–1031. doi: 10.1111/j.1467-8624.2008.01173.x
- Brannon, E. M., Lutz, D., and Cordes, S. (2006). The development of area discrimination and its implications for number representation in infancy. *Dev. Sci.* 9, F59–F64. doi: 10.1111/j.1467-7687.2006.00530.x
- Bueti, D., and Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1831–1840. doi: 10.1098/rstb.2009.0028
- Cappelletti, M., Chamberlain, R., Freeman, E. D., Kanai, R., Butterworth, B., Price, C. J., et al. (2014). Commonalities for numerical and continuous quantity skills at temporo-parietal junction. *J. Cogn. Neurosci.* 26, 986–999. doi: 10.1162/jocn_a_00546
- Cipora, K., Patro, K., and Nuerk, H.-C. (2015). Are spatial-numerical associations a cornerstone for arithmetic learning? The lack of genuine correlations suggests no. *Mind Brain Educ.* 9, 190–206. doi: 10.1111/mbe.12093
- Cleland, A. A., and Bull, R. (2015). The role of numerical and non-numerical cues in nonsymbolic number processing: evidence from the line bisection task. *Q. J. Exp. Psychol.* 68, 1844–1859. doi: 10.1080/17470218.2014.994537
- Cordes, S., and Brannon, E. M. (2009). The relative salience of discrete and continuous quantity in young infants. *Dev. Sci.* 12, 453–463. doi: 10.1111/j.1467-7687.2008.00781.x
- Crollen, V., and Noël, M.-P. (2015). Spatial and numerical processing in children with high and low visuospatial abilities. *J. Exp. Child Psychol.* 132, 84–98. doi: 10.1016/j.jecp.2014.12.006
- de Hevia, M. D. (2011). Sensitivity to number: reply to gebuis and gevers. *Cognition* 121, 253–255. doi: 10.1016/j.cognition.2011.04.002
- de Hevia, M. D., Girelli, L., and Macchi Cassia, V. (2012a). Minds without language represent number through space: origins of the mental number line. *Front. Psychol.* 3:466. doi: 10.3389/fpsyg.2012.00466
- de Hevia, M. D., Vanderslice, M., and Spelke, E. S. (2012b). Cross-dimensional mapping of number, length and brightness by preschool children. *PLoS One* 7:e35530. doi: 10.1371/journal.pone.0035530
- de Hevia, M. D., Izard, V., Coubart, A., Spelke, E. S., and Streri, A. (2014). Representations of space, time, and number in neonates. *Proc. Natl. Acad. Sci. U.S.A.* 111, 4809–4813. doi: 10.1073/pnas.1323628111
- de Hevia, M. D., and Spelke, E. S. (2009). Spontaneous mapping of number and space in adults and young children. *Cognition* 110, 198–207. doi: 10.1016/j.cognition.2008.11.003
- de Hevia, M. D., and Spelke, E. S. (2010). Number-space mapping in human infants. *Psychol. Sci.* 21, 653–660. doi: 10.1177/0956797610366091
- De Smedt, B., Noël, M.-P., Gilmore, C., and Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing skills relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends Neurosci. Educ.* 2, 48–55. doi: 10.1016/j.tine.2013.06.001

- Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., and Tsivkin, S. (1999). Sources of mathematical thinking: behavioral and brain-imaging evidence. *Science* 284, 970–974. doi: 10.1126/science.284.5416.970
- DeWind, N. K., and Brannon, E. M. (2012). Malleability of the approximate number system: effects of feedback and training. *Front. Hum. Neurosci.* 6:68. doi: 10.3389/fnhum.2012.00068
- Dormal, V., Andres, M., and Pesenti, M. (2012). Contribution of the right intraparietal sulcus to numerosity and length processing: an fMRI-guided TMS study. *Cortex* 48, 623–629. doi: 10.1016/j.cortex.2011.05.019
- Dormal, V., and Pesenti, M. (2007). Numerosity-length interference: a stroop experiment. *Exp. Psychol.* 54, 289–297. doi: 10.1027/1618-3169.54.4.289
- Dormal, V., and Pesenti, M. (2013). Processing numerosity, length and duration in a three-dimensional stroop-like task: towards a gradient of processing automaticity? *Psychol. Res.* 77, 116–127. doi: 10.1007/s00426-012-0414-3
- Ebersbach, M., Luwel, K., and Verschaffel, L. (2014). Further evidence for a spatial-numerical association in children before formal schooling. *Exp. Psychol.* 61, 323–329. doi: 10.1027/1618-3169/a000250
- Feigenson, L., Dehaene, S., and Spelke, E. (2004). Core systems of number. *Trends Cogn. Sci.* 8, 307–314. doi: 10.1016/j.tics.2004.05.002
- Fias, W., Lammertyn, J., Reynvoet, B., Dupont, P., and Orban, G. A. (2003). Parietal representation of symbolic and nonsymbolic magnitude. *J. Cogn. Neurosci.* 15, 47–56. doi: 10.1162/089989290321107819
- Gallistel, C. R. (1989). Animal cognition: the representation of space, time and number. *Annu. Rev. Psychol.* 40, 155–189. doi: 10.1146/annurev.ps.40.020189.001103
- Gebuis, T., and Gevers, W. (2011). Numerosities and space; indeed a cognitive illusion! A reply to de Hevia and Spelke (2009). *Cognition* 121, 248–252; discussion 253–255. doi: 10.1016/j.cognition.2010.09.008
- Gibson, J. J., and Gibson, E. J. (1955). Perceptual learning: differentiation or enrichment? *Psychol. Rev.* 62, 32–41. doi: 10.1037/h0048826
- Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., et al. (2015). Individual differences in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. *PLoS One* 8:e67374. doi: 10.1371/journal.pone.0067374
- Göksun, T., Woods, A. J., Chatterjee, A., Zeloni, S., Glass, L., and Smith, S. E. (2013). Elementary school children's attentional biases in physical and numerical space. *Eur. J. Dev. Psychol.* 10, 433–448. doi: 10.1080/17405629.2012.692965
- Goodwin, L. D., and Goodwin, W. L. (1999). Measurement myths and misconceptions. *Sch. Psychol. Q.* 14, 408–427. doi: 10.1037/h0089017
- Goodwin, L. D., and Leech, N. L. (2006). Understanding correlation: factors that affect the size of *r*. *J. Exp. Educ.* 74, 251–266. doi: 10.3200/JEXE.74.3.249-266
- Halberda, J., and Feigenson, L. (2008). Developmental change in the acuity of the “Number Sense”: the approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Dev. Psychol.* 44, 1457–1465. doi: 10.1037/a0012682
- Hausmann, M., Waldie, K. E., and Corballis, M. C. (2003). Developmental changes in line bisection: a result of callosal maturation? *Neuropsychology* 17, 155–160. doi: 10.1037/0894-4105.17.1.155
- Hawes, Z., Moss, J., Caswell, B., Naqvi, S., and MacKinnon, S. (2017). Enhancing children's spatial and numerical skills through a dynamic spatial approach to early geometry instruction: effects of a 32-week intervention. *Cogn. Instr.* 35, 236–264. doi: 10.1080/07370008.2017.1323902
- Hoffmann, D., Hornung, C., Martin, R., and Schiltz, C. (2013). Developing number-space associations: SNARC effects using a color discrimination task in 5-year-olds. *J. Exp. Child Psychol.* 116, 775–791. doi: 10.1016/j.jecp.2013.07.013
- Hubbard, E. M., Piazza, M., Pinel, P., and Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nat. Rev. Neurosci.* 6, 435–448. doi: 10.1038/nrn1684
- Hurewitz, F., Gelman, R., and Schnitzer, B. (2006). Sometimes area counts more than number. *Proc. Natl. Acad. Sci. U.S.A.* 103, 19599–19604. doi: 10.1073/pnas.0609485103
- Izard, V., and Spelke, E. S. (2009). Development of sensitivity to geometry in visual forms. *Hum. Evol.* 23, 213–248.
- Kaufmann, L., Vogel, S., Wood, G., Kremser, C., Schocke, M., Zimmerhackl, L., et al. (2008). A developmental fMRI study of nonsymbolic numerical and spatial processing. *Cortex* 44, 376–385. doi: 10.1016/j.cortex.2007.08.003
- Kucian, K., Grond, U., Rotzer, S., Henzi, B., Schonmann, C., Plangger, F., et al. (2011). Mental number line training in children with developmental dyscalculia. *Neuroimage* 57, 782–795. doi: 10.1016/j.neuroimage.2011.01.070
- Leibovich, T., Diesendruck, L., Rubinsten, O., and Henik, A. (2013). The importance of being relevant: modulation of magnitude representations. *Front. Psychol.* 4:369. doi: 10.3389/fpsyg.2013.00369
- Leibovich, T., and Henik, A. (2013a). Comparing performance in discrete and continuous comparison tasks. *Q. J. Exp. Psychol.* 67, 899–917. doi: 10.1080/17470218.2013.837940
- Leibovich, T., and Henik, A. (2013b). Magnitude processing in non-symbolic stimuli. *Front. Psychol.* 4:375. doi: 10.3389/fpsyg.2013.00375
- Leibovich, T., Katzin, N., Harel, M., and Henik, A. (2017). From “sense of number” to “sense of magnitude”: the role of continuous magnitudes in numerical cognition. *Behav. Brain Sci.* 40:e164. doi: 10.1017/S0140525X16000960
- Longo, M. R., and Lourenco, S. F. (2010). Bisecting the mental number line in near and far space. *Brain Cogn.* 72, 362–367. doi: 10.1016/j.bandc.2009.10.016
- Lourenco, S. F. (2015). “On the relation between numerical and non-numerical magnitudes: evidence for a general magnitude system,” in *Mathematical Cognition and Learning: Evolutionary Origins and Early Development of Basic Number Processing*, Vol. 1, eds D. C. Geary, D. B. Berch, and K. Mann Koepke (London: Elsevier Academic Press), 145–174.
- Lourenco, S. F., and Aulet, L. S. (2018). Cross-magnitude interactions across development: longitudinal evidence for a general magnitude system. *Dev. Sci.* doi: 10.1111/desc.12707 [Epub ahead of print].
- Lourenco, S. F., Ayzenberg, V., and Lyu, J. (2016). A general magnitude system in human adults: evidence from a subliminal priming paradigm. *Cortex* 81, 93–103. doi: 10.1016/j.cortex.2016.04.013
- Lourenco, S. F., and Bonny, J. W. (2017). Representations of numerical and non-numerical magnitude both contribute to mathematical competence in children. *Dev. Sci.* 20:e12418. doi: 10.1111/desc.12418
- Lourenco, S. F., Bonny, J. W., Fernandez, E. P., and Rao, S. (2012). Nonsymbolic number and cumulative area representations contribute shared and unique variance to symbolic math competence. *Proc. Natl. Acad. Sci. U.S.A.* 109, 18737–18742. doi: 10.1073/pnas.1207212109
- Lourenco, S. F., and Longo, M. R. (2010). General magnitude representation in human infants. *Psychol. Sci.* 21, 873–881. doi: 10.1177/0956797610370158
- Lourenco, S. F., and Longo, M. R. (2011). “Origins and development of generalized magnitude representation,” in *Space, Time and Number in the Brain: Searching for the Foundations of Mathematical Thought*, eds S. Dehaene and E. M. Brannon (London: Elsevier Science), 225–244. doi: 10.1016/B978-0-12-385948-8.00015-3
- McCaskey, U., von Aster, M., O’Gorman Tuura, R., and Kucian, K. (2017). Adolescents with developmental dyscalculia do not have a generalized magnitude deficit - processing of discrete and continuous magnitudes. *Front. Hum. Neurosci.* 11:102. doi: 10.3389/fnhum.2017.00102
- Moeller, K., Pixner, S., Kaufmann, L., and Nuerk, H. C. (2009). Children's early mental number line: logarithmic or decomposed linear? *J. Exp. Child Psychol.* 103, 503–515. doi: 10.1016/j.jecp.2009.02.006
- Moyer, R. S., and Landauer, T. K. (1967). Time required for judgements of numerical inequality. *Nature* 215, 1519–1520. doi: 10.1038/2151519a0
- Mussolin, C., Martin, R., and Schiltz, C. (2011). Relationships between number and space processing in adults with and without dyscalculia. *Acta Psychol.* 138, 193–203. doi: 10.1016/j.actpsy.2011.06.004
- Newcombe, N. S., Levine, S. C., and Mix, K. S. (2015). Thinking about quantity: the intertwined development of spatial and numerical cognition. *Wiley Interdiscip. Rev. Cogn. Sci.* 6, 491–505. doi: 10.1002/wcs.1369
- Nys, J., and Content, A. (2012). Judgement of discrete and continuous quantity in adults: number counts! *Q. J. Exp. Psychol.* 65, 675–690. doi: 10.1080/17470218.2011.619661
- Odic, D., Libertus, M. E., Feigenson, L., and Halberda, J. (2013). Developmental change in the acuity of approximate number and area representations. *Dev. Psychol.* 49, 1103–1112. doi: 10.1037/a0029472
- Patro, K., and Haman, M. (2012). The spatial-numerical congruity effect in preschoolers. *J. Exp. Child Psychol.* 111, 534–542. doi: 10.1016/j.jecp.2011.09.006
- Patro, K., Nuerk, H.-C., Cress, U., and Haman, M. (2014). How number-space relationships are assessed before formal schooling: a taxonomy proposal. *Front. Psychol.* 5:419. doi: 10.3389/fpsyg.2014.00419

- Pinel, P., Piazza, M., Le Bihan, D., and Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron* 41, 983–993. doi: 10.1016/S0896-6273(04)00107-2
- Price, G. R., Palmer, D., Battista, C., and Ansari, D. (2012). Nonsymbolic numerical magnitude comparison: reliability and validity of different task variants and outcome measures, and their relationship to arithmetic achievement in adults. *Acta Psychol.* 140, 50–57. doi: 10.1016/j.actpsy.2012.02.008
- Rousselle, L., Dembour, G., and Noël, M. P. (2013). Magnitude representations in Williams syndrome: differential acuity in time, space and number processing. *PLoS One* 8:e72621. doi: 10.1371/journal.pone.0072621
- Siegler, R. S., and Booth, J. L. (2004). Development of numerical estimation in young children. *Child Dev.* 75, 428–444. doi: 10.1111/j.1467-8624.2004.00684.x
- Siegler, R. S., and Opfer, J. E. (2003). The development of numerical estimation: evidence for multiple representations of numerical quantity. *Psychol. Sci.* 14, 237–243. doi: 10.1111/1467-9280.02438
- Slater, A., Mattock, A., Brown, E., and Bremner, J. G. (1991). Form perception at birth: Cohen and Younger (1984) revisited. *J. Exp. Child Psychol.* 51, 395–406. doi: 10.1016/0022-0965(91)90084-6
- Sokolowski, H. M., and Ansari, D. (2016). “Symbolic and nonsymbolic representation of number in the human parietal cortex: a review of the state-of-the-art, outstanding questions and future directions,” in *Continuous Issues in Numerical Cognition: How Many or How Much*, ed. A. Henik (San Diego, CA: Elsevier).
- Starr, A., and Brannon, E. M. (2015). Evidence against continuous variables driving numerical discrimination in infancy. *Front. Psychol.* 6:923. doi: 10.3389/fpsyg.2015.00923
- Szucs, D., Nobes, A., Devine, A., Gabriel, F. C., and Gebuis, T. (2013). Visual stimulus parameters seriously compromise the measurement of approximate number system acuity and comparative effects between adults and children. *Front. Psychol.* 4:444. doi: 10.3389/fpsyg.2013.00444
- Tudusciuc, O., and Nieder, A. (2007). Neuronal population coding of continuous and discrete quantity in the primate posterior parietal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 104, 14513–14518. doi: 10.1073/pnas.0705495104
- Van Opstal, F., and Verguts, T. (2013). Is there a generalized magnitude system in the brain? Behavioral, neuroimaging, and computational evidence. *Front. Psychol.* 4:435. doi: 10.3389/fpsyg.2013.00435
- van Vugt, P., Fransen, I., Creten, W., and Paquier, P. (2000). Line bisection performances of 650 normal children. *Neuropsychologia* 38, 886–895. doi: 10.1016/S0028-3932(99)00130-X
- Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends Cogn. Sci.* 7, 483–488. doi: 10.1016/j.tics.2003.09.002
- WMA (2002). *The World Medical Association's Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects*. Washington, DC: WMA General Assembly.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Kucian, McCaskey, von Aster and O’Gorman Tuura. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Commentary : The Developmental Trajectory of the Operational Momentum Effect

Martin H. Fischer^{1*}, Alex Miklashevsky¹ and Samuel Shaki²

¹ Psychology Department, University of Potsdam, Potsdam, Germany, ² The Department of Behavioral Sciences, Ariel University, Ariel, Israel

Keywords: embodied cognition, operational momentum, SNARC effect, mental arithmetic, numerical cognition

A Commentary on:

The Developmental Trajectory of the Operational Momentum Effect

by Pinheiro-Chagas, P., Didino, D., Haase, V. G., Wood, G., and Knops, A. (2018). *Front. Psychol.* 9:1062. doi: 10.3389/fpsyg.2018.01062

OPEN ACCESS

Edited by:

Krzysztof Cipora,
Universität Tübingen, Germany

Reviewed by:

Catherine Thevenot,
Université de Lausanne, Switzerland
Jérôme Prado,
Centre national de la recherche
scientifique (CNRS), France

*Correspondence:

Martin H. Fischer
martinf@uni-potsdam.de

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 27 August 2018

Accepted: 30 October 2018

Published: 21 November 2018

Citation:

Fischer MH, Miklashevsky A and
Shaki S (2018) Commentary : The
Developmental Trajectory of the
Operational Momentum Effect.
Front. Psychol. 9:2259.
doi: 10.3389/fpsyg.2018.02259

Recently, Pinheiro-Chagas et al. (2018) studied the development of the operational momentum (OM) which denotes a tendency to accept larger than correct outcomes in addition and smaller than correct outcomes in subtraction. The authors reviewed some theories of OM and derived two competing predictions. First, they described the attentional account, according to which OM results from an overshoot of an attentional spotlight when moving along the spatially oriented mental number line (MNL) in accordance with the magnitude of the second operand. Given that “formal schooling ... might consolidate a systematic movement direction during the acquisition of arithmetical skills” (p.3), older children should show more OM. Secondly, they described the compression account of OM according to which linear operations (addition, subtraction) are performed on logarithmically compressed operand representations. Referring to a log-to-linear developmental shift in the placement of numbers on visually presented number lines, they predicted that older children should show less (un-) compression and thus less OM. Their results from 8 to 12-year olds showed a gradual increase of OM starting at 9 years and thus supported the attentional account.

The clear performance pattern reported by Pinheiro-Chagas and colleagues makes a useful contribution to the literature on OM development but their report also misrepresents the state of knowledge about OM. It might leave readers unnecessarily misinformed about the multi-faceted origin of this bias generally, and more specifically about the status of reverse OM for our understanding of cognitive biases in formal reasoning. We draw attention to these points below.

First, the authors acknowledged early OM in 9-month-olds (McCrink and Wynn, 2009) as well as reverse OM in 6-year-olds (Knops et al., 2013), thus recognizing a potential problem with their conclusion of late-emerging and gradually increasing OM. While the authors mentioned the work of Pinhas and colleagues they did not convey its full impact with regard to this point. First, Pinhas and Fischer (2008; see also Shaki et al., 2018) observed larger OM in zero problems (such as 4+0) compared to non-zero problems (e.g., 3+1). This alone could suffice to discredit the compression account because the logarithm of zero is not defined. Thus, the compression account was arguably a mere strawman pitted against the attentional account, although other methodological differences, such as the number format, remain. But if attention shift magnitude is “... a distance corresponding to the magnitude of the second operand” (p. 2), how does this account explain larger OM with zero problems?

Further inconsistencies are reflected in the methods: From an attentional perspective, repeated downward movements of both addends, as well as upward movements of the subtrahend,

constitute inconsistencies with the vertical MNL that maps small quantities below larger quantities. Experience with vertical mappings will change over age and might increase the performance consequences of such inconsistencies. More generally, why were operations along a horizontal MNL primed with vertical movements? The fact that subtrahends moved away from the area of interest in the display center removed attention from the place of mentally simulating the outcome, thus impeding subtraction.

Secondly, Pinhas and Fischer (2008) proposed multiple sources of OM, including the operands, the operator, and the result. Taking into consideration evidence from biased quantitative reasoning, estimation heuristics and spatial-numerical associations, we have since developed this proposal into a comprehensive model of arithmetic heuristics and biases (AHAB; see also Shaki and Fischer, 2017; Fischer and Shaki, 2018; Shaki et al., 2018). This model can explain the complete range of findings reported in the literature, including reverse OM, as a weighted contribution from an anchoring effect, an estimation heuristic, and spatial associations of operands and operators. Pinheiro-Chagas et al.'s report created the false impression that reverse OM is an anomaly. Instead it was found repeatedly (Charras et al., 2012, 2014; Knops et al., 2013; Pinhas et al., 2015; Blini et al., 2018) and can be understood as reflecting anchoring bias in non-zero problems. However, anchoring bias increases from fourth to eight grade (Smith, 1999) and this should gradually reduce OM unless other factors compensate for this bias.

One further strength of AHAB is its ability to account for both spatial and non-spatial biases in mental arithmetic, regardless of whether computational uncertainty originated from encoding non-symbolic operators or results, as in studies by Knops and colleagues, or from mapping of perfectly identifiable operators and results onto a continuous response dimension, such as

horizontal lines or time intervals (Shaki et al., 2015). It would be interesting to learn whether Pinheiro-Chagas and colleagues replicated the spatial bias in response selection previously observed in this paradigm by Knops et al. (2009).

Finally, the authors mention also an heuristic account of OM: a tendency to accept more than the correct outcome for additions and less than the correct outcome for subtractions because addition leads to “more” and subtraction to “less” (McCrink and Wynn, 2009). They compare it to the attentional account and state that “... the two accounts provide equivalent predictions” (p. 3). This is in conflict with the recent analysis offered in McCrink and Hubbard (2018, p. 240) that “... the use of heuristics is generally increased when attention is decreased”. We think that heuristics are triggered by operators. Yet, OM only emerges late, i.e., when both operator and second operand have been processed (Liu et al., 2017; Masson et al., 2017; Blini et al., 2018). Results obtained from procedures where operators even preceded the first operand (cf. Knops et al., 2009) or multiple quantities are presented during responding (cf. Pinheiro-Chagas et al., 2018) must be interpreted cautiously because the normal ingredients of OM are dis-ordered or diluted.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

Supported by DFG FI 1915/8-1 Competing heuristics and biases in mental arithmetic. We acknowledge the support of the Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of University of Potsdam.

REFERENCES

- Blini, E., Pitteri, M., and Zorzi, M. (2018). Spatial grounding of symbolic arithmetic: an investigation with optokinetic stimulation. *Psychol. Res.* 1–20. doi: 10.1007/s00426-018-1053-0
- Charras, P., Brod, G., and Lupianez, J. (2012). Is $26 + 26$ smaller than $24 + 28$? Estimating the approximate magnitude of repeated vs. different numbers. *Attention Percept. Psychophys.* 74, 163–173. doi: 10.3758/s13414-011-0217-4
- Charras, P., Molina, E., and Lupianez, J. (2014). Additions are biased by operands: evidence from repeated versus different operands. *Psychol. Res.* 78, 248–265. doi: 10.1007/s00426-013-0491-y
- Fischer, M. H., and Shaki, S. (2018). Number concepts: abstract and embodied. *Phil. Trans. R. Soc. B* 373:20170125. doi: 10.1098/rstb.2017.0125
- Knops, A., Viarouge, A., and Dehaene, S. (2009). Dynamic representations underlying symbolic and nonsymbolic calculation: evidence from the operational momentum effect. *Attention Percept. Psychophys.* 71, 803–821. doi: 10.3758/APP.71.4.803
- Knops, A., Zitzmann, S., and McCrink, K. (2013). Examining the presence and determinants of operational momentum in childhood. *Front. Psychol.* 4:325. doi: 10.3389/fpsyg.2013.00325
- Liu, D., Cai, D., Verguts, T., and Chen, Q. (2017). The time course of spatial attention shifts in elementary arithmetic. *Sci. Rep.* 7:921. doi: 10.1038/s41598-017-01037-3
- Masson, N., Letesson, C., and Pesenti, M. (2017). Time course of overt attentional shifts in mental arithmetic: evidence from gaze metrics. *Quart. J. Exp. Psychol.* 71, 1009–1019. doi: 10.1080/17470218.2017.1318931
- McCrink, K., and Hubbard, T. (2018). Dividing attention increases operational momentum. *J. Numerical Cogn.* 382, 230–245. doi: 10.5964/jnc.v3.i2.34
- McCrink, K., and Wynn, K. (2009). Operational momentum in large-number addition and subtraction by 9-month-olds. *J. Exp. Child Psychol.* 103, 400–408. doi: 10.1016/j.jecp.2009.01.013
- Pinhas, M., and Fischer, M. H. (2008). Mental movements without magnitude? A study of spatial biases in symbolic arithmetic. *Cognition* 109, 408–415. doi: 10.1016/j.cognition.2008.09.003
- Pinhas, M., Shaki, S., and Fischer, M. H. (2015). Addition goes where the big numbers are: evidence for a reversed operational momentum effect. *Psychonomic Bull. Rev.* 22, 993–1000. doi: 10.3758/s13423-014-0786-z
- Pinheiro-Chagas, P., Didino, D., Haase, V. G., Wood, G., and Knops, A. (2018). The developmental trajectory of the operational momentum effect. *Front. Psychol.* 9:1062. doi: 10.3389/fpsyg.2018.01062

- Shaki, S., and Fischer, M. H. (2017). Competing biases in mental arithmetic: when division is more and multiplication is less. *Front. Hum. Neurosci.* 11:37. doi: 10.3389/fnhum.2017.00037
- Shaki, S., Pinhas, M., and Fischer, M. H. (2018). Heuristics and biases in mental arithmetic: revisiting and reversing operational momentum. *Think. Reason.* 24, 138–156. doi: 10.1080/13546783.2017.1348987
- Shaki, S., Sery, N., and Fischer, M. H. (2015). 1 + 2 is more than 2 + 1: violations of commutativity and identity axioms in mental arithmetic. *J. Cogn. Psychol.* 27:4. doi: 10.1080/20445911.2014.973414
- Smith, H. D. (1999). Use of anchoring and adjustment heuristic by children. *Curr. Psychol.* 18, 294–300.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Fischer, Miklashevsky and Shaki. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Linear Spatial–Numeric Associations Aid Memory for Single Numbers

John Opfer^{1*}, Dan Kim¹, Christopher J. Young² and Francesca Marciani³

¹ Department of Psychology, The Ohio State University, Columbus, OH, United States, ² University of Chicago Consortium, Chicago, IL, United States, ³ University of Alabama at Birmingham, Birmingham, AL, United States

Memory for numbers improves with age. One source of this improvement may be learning linear spatial–numeric associations, but previous evidence for this hypothesis likely confounded memory span with quality of numerical magnitude representations and failed to distinguish spatial–numeric mappings from other numeric abilities, such as counting or number word–cardinality mapping. To obviate the influence of memory span on numerical memory, we examined 39 3- to 5-year-olds' ability to recall one spontaneously produced number (1–20) after a delay, and the relation between numeric recall (controlling for non-numeric recall) and quality of mapping between symbolic and non-symbolic quantities using number-line estimation, give-a-number estimation, and counting tasks. Consistent with previous reports, mapping of numerals to space, to discrete quantities, and to numbers in memory displayed a logarithmic-to-linear shift. Also, linearity of spatial–numeric mapping correlated strongly with multiple measures of numeric recall (percent correct and percent absolute error), even when controlling for age and non-numeric memory. Results suggest that linear spatial–numeric mappings may aid memory for number over and above children's other numeric skills.

Keywords: numerical estimation, memory development, numerical cognition, spatial–numerical association, memory, counting, cardinality knowledge

OPEN ACCESS

Edited by:

Hans-Christoph Nuerk,
University of Tübingen, Germany

Reviewed by:

Koen Luwel,
KU Leuven, Belgium
Christine Schiltz,
University of Luxembourg,
Luxembourg

*Correspondence:

John Opfer
opfer.7@osu.edu

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 20 December 2017

Accepted: 16 January 2019

Published: 04 February 2019

Citation:

Opfer J, Kim D, Young CJ and
Marciani F (2019) Linear
Spatial–Numeric Associations Aid
Memory for Single Numbers.
Front. Psychol. 10:146.
doi: 10.3389/fpsyg.2019.00146

INTRODUCTION

Both in school and everyday life, children are presented with a potentially dazzling succession of numbers to remember. Some numbers must be remembered exactly, such as phone numbers and the answers to arithmetic problems. Others only need to be remembered approximately, such as the number of children in one's class, the amount of money in one's piggy bank, or the temperature forecast for tomorrow's weather. When confronted with a series of numbers in either type of situation—e.g., a digit span task (Dempster, 1981) or a vignette (Brainerd and Gordon, 1994)—young children recall numbers much less accurately than older children and adults. In this paper, we examine whether developmental changes in numerical representation accounts for individual differences in memory for numbers. Specifically, we test how children's memory for numbers relates to their memory for non-numeric information (e.g., color) and to their knowledge of numeric magnitude, indexed by their ability to map a number to a spatial location on a number line, to map a discrete number of objects to a number word, and to count.

Fuzzy Trace Theory

The fuzzy trace theory (FTT) depicts information as being stored in memory using one of two representational formats, a short-term verbatim or “surface form” representation and a long-term

gist or “fuzzy trace” representation (Reyna et al., 2009). Within this account, when numerical information is learned (e.g., Farmer Brown owns many animals. He has 5 sheep, 11 cows, and 3 dogs.), numbers can be encoded precisely using a verbatim representation (e.g., 5 sheep, 11 cows, and 3 dogs), including the specific format in which they were presented (e.g., numerals, dots, etc.). Verbatim memories preserve exact surface forms of numerical inputs for a short period of time, but lack relative concepts or relations between numbers (e.g., *the most*, *the least*, *more*, and *less*). However, the meaning of numbers is encoded as a gist representation (e.g., Farmer Brown owns more cows than dogs.). Gist memories do not contain formatting information of numerical inputs, but preserve a sense of approximate magnitude or relative amount (e.g., *about six*, *less*, *more*, *a lot*, etc.) for a longer period of time (Brainerd and Gordon, 1994).

An attractive feature of FTT is that it helps to explain what changes in memory development. That is, for young children, memory for numerical information in the verbatim representation is superior to that of the gist representation, but this advantage attenuates with age. Thus, by adulthood, there is greater reliance on longer lasting, but imprecise, gist representations of numerical magnitude (Brainerd and Gordon, 1994). Empirical support for this account was shown by dissociation between performance on relative comparisons of quantity in gist tests (e.g., “Which of Farmer Brown’s animals are the most, cows or dogs?”) and exact identification of quantity in verbatim memory tests (e.g., “How many cows does Farmer Brown own, 11 or 3?”). Between preschool and second grade, gist recall increased with age and ultimately surpassed verbatim recall, which did not change with age.

While FTT provides accurate predictions for general improvement via an age-related switch to gist memory for numbers, it is not clear whether it provides a sufficient mechanism for how features of the stimulus, such as the magnitude of a to-be-remembered number, will influence the likelihood of its recall. One idea had been that “physical distinctiveness” helps to distinguish items in the verbatim representation and thereby improve memory (Brainerd and Reyna, 1993; Brainerd and Gordon, 1994, p. 166). Under this view, the physical distinctiveness of some items in verbatim tests is greater than that of others. For example, when choosing whether 12 vs. 10 cows had been studied, the physical difference between the two two-digit numbers is less than the physical difference between 12 vs. 3 cows. Consistent with this idea, greater numerical distance improves young children’s memory more than it does older children’s memory.

Representational Change Account

A somewhat different depiction of the distance effect—and how symbolic numeric information is stored in memory—comes from findings on development of numerical magnitude representations (for review, see Opfer and Siegler, 2012). As we will see, this account also depicts representations of numeric magnitude as being “fuzzy” and approximate. However, unlike FTT, the internal magnitudes associated with numbers are also depicted as changing with development, such that the

distinctiveness of the memory trace changes more for large than small numbers.

A coherent picture of how numerical magnitude representations change with age and experience is provided by previous research on numerical estimation. In early development, young preschoolers attach no cardinal value to numerical symbols, and they do not yet map numeric symbols like number words and Arabic numerals (even approximately) to non-symbolic numeric quantities. For example, 2- and 3-year-olds who count flawlessly from 1 to 10 have no idea that $6 > 4$, nor do preschoolers of these ages know how many objects to give an adult who asks for 4 or more (Le Corre et al., 2006; though see Lee and Sarnecka, 2010). Later, as non-mapping children gain experience associating numeric symbols with real-world quantities (such as sets of objects or number of sounds), they initially map numeric symbols to a noisy, logarithmically compressed mental number line which represents and stores the magnitudes of non-symbolic quantities (such as objects and tones) in memory. During this period, preschoolers know approximately how many objects to give an adult who asks for 1–20 objects and approximately where on the number-line 1–20 fall, but their estimates in each case increase logarithmically with the number to be estimated (Berteletti et al., 2010; Opfer et al., 2010; Thompson and Siegler, 2010; Kim and Opfer, 2017).

Development from a logarithmic to linear representation of numeric value occurs iteratively. Over a period that typically lasts 1–3 years for a given numerical range (0–10, 0–20, 0–100, or 0–1,000), children’s mapping of symbolic numbers to non-symbolic quantities changes from a logarithmically compressed form to a linear form, where subjective and objective numerical values increase in a 1:1 fashion (Siegler and Opfer, 2003; Siegler and Booth, 2004; Opfer and Thompson, 2008; Thompson and Opfer, 2008, 2010; Berteletti et al., 2010; Opfer et al., 2010). Use of linear numerical-magnitude representations occurs earliest for the numerals that are most frequent in the environment (i.e., the smallest whole numbers) and is extended to increasingly large numbers with experience (Thompson and Opfer, 2010). Although some alternative models (e.g., Gallistel and Gelman, 1992; Ebersbach et al., 2008; Cantlon et al., 2009; Moeller et al., 2009; Barth and Paladino, 2011; Cohen and Sarnecka, 2014) have argued that the mapping of symbolic numbers to magnitudes does not show an abrupt transition from a precisely logarithmic to a precisely linear representation (but see Opfer et al., 2011, 2016; Young and Opfer, 2011; Kim and Opfer, 2017; Qin et al., 2017), all models capture a similar phenomenon—young children estimate the magnitudes of small numbers as differing more than the magnitudes of large numbers, whereas older children and adults estimate the magnitudes in a more closely 1:1 fashion.

Associations between linear numerical-magnitude representations and numeric memory were recently explored by Thompson and Siegler (2010), who presented children with numbers in a vignette and asked them to recall the numbers after a brief distracter task (naming four cartoon characters). For example, children were given a story, “Colleen washes the dishes at a restaurant. This month, she washed N_1 forks, N_2 cups, and N_3 plates.” After a distracter task, they were asked how many forks, cups, and plates Colleen washed respectively. Several

observations from Thompson and Siegler (2010) suggested that linear numerical-magnitude representations aided memory for numerical information. First, linearity and accuracy on approximate magnitude tasks (number-line estimation and number categorization) were highly correlated with number memory, whereas accuracy on non-approximate magnitude tasks (counting and number identification) was not. Thus, a third variable (such as overall numeric proficiency) was unlikely to be a source of the positive correlation between numerical estimation and memory. Second, memory accuracy measured using the percent absolute error (PAE) deteriorated with the magnitude of the number given, especially for children with a logarithmic representation of number (Thompson and Siegler, 2010, Experiment 3). This finding is important because if numeric symbols are mapped with a constant noisiness to a logarithmically scaled mental number line, then signal overlap increases dramatically with numerical value, thereby leading to significant interference from adjacent values as the target number increases. Interference from highly similar exemplars is a well-known source of errors in recall (Schacter et al., 1998), yet it would not be predicted if children's memory for numbers depended solely on their memory span. Finally, preschoolers' difficulty recalling large numbers could not be explained by large numbers simply being unfamiliar to them. When preschoolers were tested to see how high they could count, Thompson and Siegler (2010) observed no correlation between the largest number counted and memory accuracy.

The Current Study

In this study, we investigated a potential source of concern about evidence supporting the representational change account. That is, individual and developmental differences also exist in memory span; the number of digits that can be accurately recalled at age 2 years is about 2, at age 5 about 4, at age 10 about 5, and among adults 7 ± 2 (Dempster, 1981). Thus, given that memory span and linear numerical-magnitude representations (in the 1–20 range) develop simultaneously, a spurious correlation between linearity of numerical estimation and span-based numerical memory may have been observed by Thompson and Siegler (2010) because children were asked to remember *multiple* items that exceeded their memory span.

This concern seems particularly justified by two previous findings. First, a correlation exists between working memory span and linear numerical-magnitude representations (e.g., Geary et al., 2007). Second, the sum of items and distractors in Thompson and Siegler (2010) study would have been at the edge of many children's memory span, leading many children to fail to recall numeric information if memory span were a contributor to numerical memory.

To address these concerns, the current study tested 3- to 5-year-olds' memory for a *single number*, thereby obviating any potential contribution of individual differences in memory span to numerical memory, and children's memory for a *single color*, in order to control for non-numeric memory ability. As in Thompson and Siegler (2010; Experiment 2), we examined (1) preschoolers' recall of numbers 1–20 because preschoolers vary in whether they represent these numbers as increasing linearly

(Berteletti et al., 2010; Thompson and Siegler, 2010), (2) the degree to which preschoolers' estimates of positions of numbers on number lines increased linearly with actual numeric value, and (3) preschoolers' counting from 1 to 20. Additionally, we examined (4) children's performance on a "give-a-number" task (Wynn, 1990) because estimates on this task have been reported to show a logarithmic-to-linear shift (Opfer et al., 2010) and to provide a more robust test of number understanding than counting accuracy (Wynn, 1990; Le Corre et al., 2006; Sarnecka and Carey, 2008).

Our predictions regarding quantitative performance were derived from research on development of numerical abilities. Generally, we predicted development across all three tasks would ultimately involve accurate and linear mappings between symbols and quantities, but the developmental paths to this would be more similar for number-line and give-a-number estimation than counting. This is because accurately translating from numerical magnitudes to symbolic numbers can be accomplished procedurally (by counting), without knowing the 1:1 mapping of symbols and quantities more generally (Briars and Siegler, 1984; Wynn, 1990). In contrast, accurately translating from a symbolic number to a magnitude requires this knowledge, and it develops slowly from non-mapping (i.e., random translation between symbols and magnitudes) to noisy non-linear mapping to precise linear mapping (i.e., systematic translation between symbols and magnitudes) (Siegler and Opfer, 2003; Opfer et al., 2010; Wagner and Johnson, 2011). Thus, among children who map symbols to quantity in the number-line estimation and/or the give-a-number task, we expected a significant increase in linearity with age along with an increase in accuracy. In contrast, among non-mappers, we expected no age-related improvements in linearity or accuracy.

Our predictions regarding memory performance were derived from the representational change account. Specifically, if linearity of numeric magnitude representations influences the likelihood that numbers are recalled accurately, then preschoolers with the most linear mappings on our estimation tasks would likely recall numbers the most accurately as well. Further, if representations of numeric magnitude develop from a logarithmic (or similarly compressed) mapping to a linear mapping over preschool, an interaction between age and magnitude on memory accuracy would be expected. That is, young and old preschoolers would have nearly equally accurate memories for small numbers, whereas young preschoolers would have significantly less accurate memories for large numbers than older preschoolers. If so, linearity of representation would likely mediate the relations between age and number memory that increases with age.

MATERIALS AND METHODS

Participants

Thirty-nine preschoolers (54% female) were recruited from six child-care centers in the Columbus metro area. Preschoolers were aged 3 years ($n = 13$, $M = 3.63$), 4 years ($n = 14$, $M = 4.45$), and 5 years ($n = 12$, $M = 5.35$). Mean age of all children was 4.45 years. Only children whose parents or guardians had provided written

consent and who verbally agreed to take part in the research participated in the study. All task materials and experimental procedures described below were approved by the Institutional Review Board at The Ohio State University.

Tasks

For all tasks, preschoolers were presented with eight numbers in randomized order. We presented the same numbers used in Berteletti et al. (2010): 2, 4, 6, 7, 13, 15, 16, and 18 to each child on each task.

Number-Line Estimation

The number-line estimation task was adapted from Siegler and Opfer (2003). Preschoolers were presented with a sheet of paper on each trial of the task. Centered on each sheet was a 25-cm line, flanked by two vertical hatch marks. The value “0” was written below the vertical hatch mark representing the left end of the line, and the value “20” was written below the mark representing the right end of the line. Above the middle of the line was one of the 8 task numerals, centered within a circle. The experimenter told the child, “Today, we’re going to play a game with number lines. What I’m going to ask you to do is show me where on the number line some numbers are. When you decide where the number goes, I want you to make a mark through the number line like this,” and demonstrated marking the line. All numbers were read aloud, but preschoolers were not corrected on their responses nor told that the halfway position along the line is where “10” should go.

Give-a-Number Estimation

The give-a-number estimation task was adopted from Wynn (1990). Preschoolers were presented with a pile of 20 blue poker chips and told that the experimenter would ask them for a number of chips. The child’s task was to place what he or she believed to be the correct number of chips before the experimenter, and the experimenter confirmed the child’s response by asking, “And how many is that?”

Counting Task

In the counting task, preschoolers were presented a 72-cm black poster board strip. Attached to each strip were a number of white poker chips that were presented to each child so that the first chip was at the left end of the strip, with each successive chip centered 4 cm to the right of the previous one. Each child was told, “You have to find out how many chips there are on this card.” Children were neither encouraged nor discouraged to count, so that they would use their own strategies. Thus, although it was possible for children to estimate the number of chips, most children of all ages counted chips aloud from left to right.

Number/Color Recall Task

The numerical recall task was intertwined with the counting task. A recall trial immediately followed each counting trial. After explaining the counting task to the child, the experimenter indicated a second experimenter at a separate location within the same room. The experimenter then instructed the child that he or she was to tell the second experimenter a “password” and how many chips there were on the card. The “password,” designed to

prevent children from rehearsing the number prior to recall, was the color of construction paper presented by the experimenter. Upon reaching the second experimenter, the child was asked the “password” (color), and then how many chips were on the card (number). Thus, by testing recall of numbers and colors that children had generated themselves, we could be certain that the items to be remembered were familiar to children and had been encoded.

Design and Procedure

Testing was administered in two sessions. In a first session, children played one of two games based on Siegler and Ramani (2009). In each game, 22 colored squares of identical size were ordered consecutively on a board. The first square was labeled “Start,” and the last square was labeled “Finish.” Squares between the first and last were consecutively numbered from 1 to 20. The sole difference between the games was the arrangement of numbers. In one game, numbered squares were placed in a horizontal line across the board, arranged left-to-right. In the other game, numbered squares were arranged in a circle, with numbers increasing in value in a clockwise direction. In the games, children were asked to move their token from “Start” to “Finish” and read the numbers on the squares as they moved. The games were included to test effects of the spatial arrangement of numbers on children’s numerical understanding. Unlike Siegler and Ramani (2009), however, there were no main (or interactive) effects of game type on number and memory tasks for the second session, presumably due to the much shorter time allotted for game play in the present study.

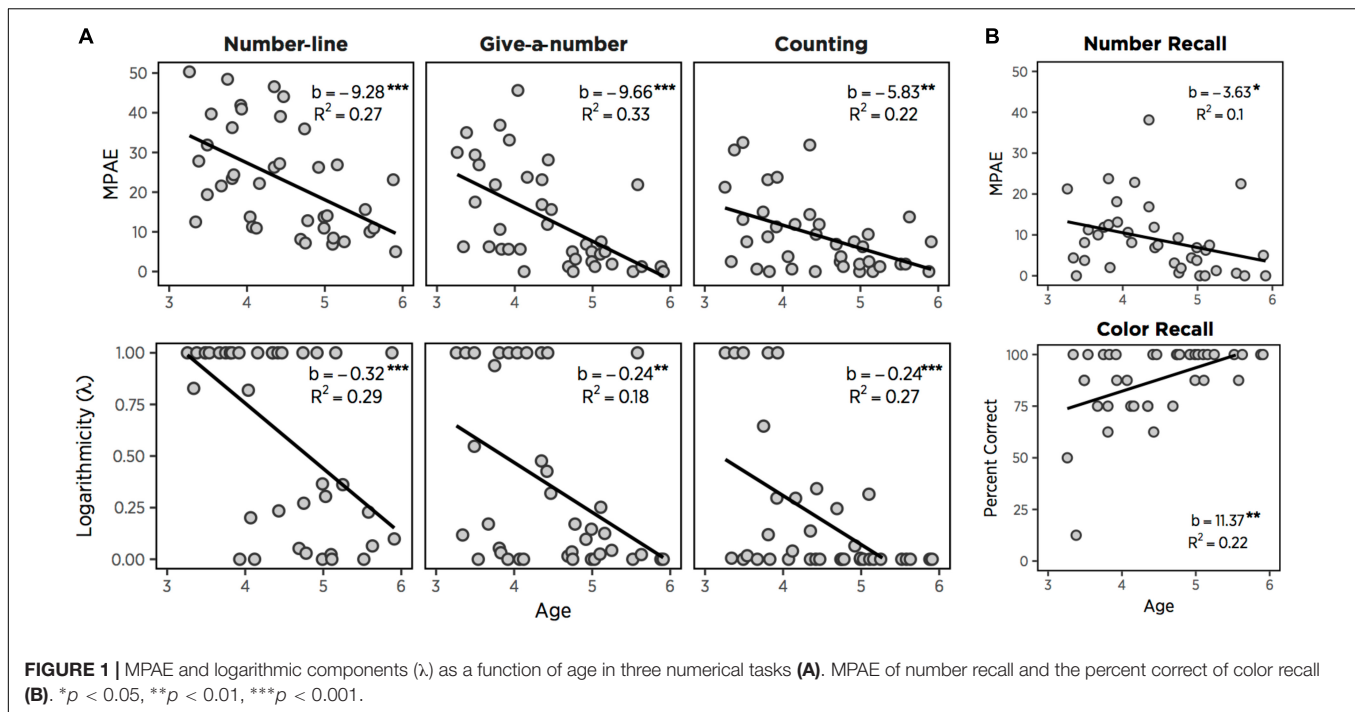
In a second session, experimenters revisited schools within 4 days to administer the battery of tasks described above (i.e., number-line, give-a-number, counting, and recall tasks). Order of presentation was counterbalanced, with the exception that the recall task necessarily followed the counting task. There were no carry-over effects of particular task order ($ps > 0.05$). Children were tested individually during one 25-min session occurring in a quiet room in their school.

RESULTS

Our results are divided into two major sections. In the first section (“Description of Task Performance”), we describe age-related changes in preschoolers’ number-line and give-a-number estimation, counting, and recall. In the next section (“Logarithmic Compression in Numerical Tasks as Predictors of Memory Performance”), we examine relations among quantitative performance and recall. One 4-year-old who completed number-line and give-a-number tasks but did not complete counting and recall tasks was excluded from analyses that involved the two incomplete tasks.

Description of Task Performance

For our quantitative tasks, we examined accuracy and linearity of the mapping between numeric symbols and quantities. Accuracy was measured using the mean percent absolute error



(MPAE) scores for a child. Within each trial, PAE for number-line estimation was calculated using the formula, $(|\text{Number Presented} - \text{Number Estimated}|/20) \times 100$, for give-a-number estimation using $(|\text{Chips Requested} - \text{Chips Given}|/20) \times 100$, and for counting using $(|\text{Chips Shown} - \text{Number Counted}|/20) \times 100$. The MPAE was then computed by obtaining the across-trial mean of the PAEs.

To calculate linearity, children's responses of the three tasks were fitted by a mixed log-linear model (MLLM) (Anobile et al., 2012; Opfer et al., 2016), formalized as follows:

$$y = a \left(\lambda \frac{U}{\ln(U)} \ln(x) + (1 - \lambda)x \right).$$

In the MLLM, a denotes a scaling parameter, U the upper bound, x a given magnitude, and y a child's estimate. It also includes a logarithmic component (λ) that measures a degree of logarithmic compression in responses and is constrained to be between 0 and 1. If estimates are perfectly logarithmic, λ equals 1, whereas λ is 0 for perfectly correct and linear estimates. The MLLM was fitted to each child individually and to median responses collapsed across children by age group.

Number-Line Estimation

As expected, estimation accuracy (MPAE) improved significantly with age, $b = -9.281$, $t(37) = -3.67$, $p < 0.001$. Also, linearity measured with logarithmic components (λ) improved with age, $b = -0.316$, $t(37) = -3.90$, $p < 0.001$ (Figure 1A). Age also explained 26.6% of variance in accuracy, $F(1,37) = 13.44$, $p < 0.001$, and 29.1% of variance in linearity, $F(1,37) = 15.19$, $p < 0.001$. The average MPAE for all children was 23.16 ($SD = 13.47$), and the averaged value of logarithmic components was 0.61 ($SD = 0.44$) (Table 1). The MPAE was 32.19 ($SD = 11.75$)

for 3-year-olds, 23.68 ($SD = 13.59$) for 4-year-olds, and 12.76 ($SD = 6.56$) for 5-year-olds. The average logarithmic component (λ) for 3-year-olds was 0.91 ($SD = 0.28$), 0.61 ($SD = 0.44$) for 4-year-olds, and 0.29 ($SD = 0.36$) for 5-year-olds. These results are broadly consistent with the "logarithmic-to-linear shift" in number-line estimation (Siegler et al., 2009).

Previous work explained age-related changes in accuracy of preschoolers' number-line estimates as coming from a logarithmic to linear shift in representations of numerical magnitude (Berteletti et al., 2010; Thompson and Siegler, 2010). Consistent with this idea, the accuracy measure (MPAE) significantly correlated with logarithmic components in number-line estimation $r(36) = 0.64$, $p < 0.001$ (Table 2). The association remained strongly even after controlling for age, $r(35) = 0.50$, $p < 0.01$. Besides, we found that linearity of estimates improved with age. Median estimates of 3-year-olds increased more logarithmically with actual value, whereas estimates of 4- and 5-year-olds increased more linearly with actual value compared to younger children (3s: $\lambda = 1$, 4s: $\lambda = 0.34$, 5s: $\lambda = 0.14$). Thus, at the group level, all three age groups mapped numeric magnitudes at least approximately to the number line, with logarithmic compression decreasing with age.

To test whether individual children approximately mapped the magnitude of symbolic numbers to their number-line estimates as well, we next evaluated whether each child's estimates increased with the numbers presented to them. To do this, we categorized children into two categories, mapping and non-mapping groups, by using a goodness of fit measure (R^2) of the MLLM. Children whose estimates did not increase progressively with given magnitudes and were not explained by a MLLM at all (i.e., $R^2 = 0$) were considered as non-mappers, whereas children whose estimates were accounted for by a MLLM (i.e., $R^2 > 0$)

regardless of the statistical significance of R^2) were categorized as mappers (but see Sella et al., 2017, for mapper categorization using a simple linear or log function). The non-mappers ($n = 17$), constituted 43.6% of all preschoolers (69.2% of 3-year-olds, 42.9% of 4-year-olds, and 16.7% of 5-year-olds). These results indicate a significant difficulty among the majority of children, particularly 3- and 4-year-olds, in mapping symbolic numbers even approximately to their corresponding positions on the number-line.

Among preschoolers who did not show this difficulty in mapping numbers to the number line ($n = 22$), however, there was stronger support for a logarithmic-to-linear shift. First, we observed significant age-related improvements in linearity, $b = -0.319$, $t(20) = -3.88$, $p < 0.001$, and in accuracy, $b = -6.417$, $t(20) = -2.57$, $p < 0.05$. Age explained 42.9% of variance in linearity, $F(1,20) = 15.05$, $p < 0.001$, and 24.7% of variance in accuracy, $F(1,20) = 6.59$, $p < 0.05$. Consistent with the individual level analyses, median estimates of mappers were more linear with age. As shown in **Figure 2**, median estimates by 3-year-olds were logarithmically compressed ($\lambda = 0.73$), whereas 4- and 5-year-olds produced linear median estimates (4s: $\lambda = 0.09$, 5s: 0.07). In contrast, non-mapping children ($n = 17$) showed no effect of age on either linearity or accuracy ($ps > 0.05$). These results suggest that while mapping children show a developmental progression in the linearity of the representation used on the number line estimation task as well as better accuracy, non-mapping children seem to have had such a poor understanding of the mapping of numeric magnitudes to linear distance that neither their accuracy nor their linearity improved with age¹.

¹One may argue that the MLLM is not the best model for young children's estimates, and their number-line performance may be better explained by a variety of cyclic power models (Barth and Paladino, 2011; Slusser et al., 2013; Cohen and Sarnecka, 2014; Cohen and Quinlan, 2018). To test this, we also fitted the mixed cyclic-power models (MCPM1 and MCPM2; see Kim and Opfer, 2017, for details), the mixtures of diverse cyclic-power models that have been proposed in literature (Opfer et al., 2016; Kim and Opfer, 2017; Qin et al., 2017). When the MCPMs were compared with the MLLM using corrected Akaike information criterion (AICc) and Bayesian information criterion (BIC), the MLLM was the best-fitting model for median estimates of all three age groups and all individual children's estimates regardless of which model-selection criterion was used. Furthermore, the weight

Give-a-Number Estimation

Give-a-number estimation accuracy also improved significantly with age, $b = -9.662$, $t(37) = -4.25$, $p < 0.001$, as did linearity, $b = -0.240$, $t(37) = -2.86$, $p < 0.01$ (**Figure 1A**). Age explained 32.8% of variance in accuracy, $F(1,37) = 18.03$, $p < 0.001$, and 18.1% of variance in linearity, $F(1,37) = 8.16$, $p < 0.01$. The average MPAE for all children was 12.92 ($SD = 12.64$), and average logarithmic component was 0.36 ($SD = 0.42$) (**Table 1**). The MPAE was 20.38 ($SD = 12.26$) for 3-year-olds, 13.35 ($SD = 13.19$) for 4-year-olds, and 4.32 ($SD = 5.99$) for 5-year-olds. The average logarithmic component for 3-year-olds was 0.53 ($SD = 0.47$), 0.40 ($SD = 0.43$) for 4-year-olds, and 0.13 ($SD = 0.28$) for 5-year-olds.

Previous work had explained age-related changes in accuracy of preschoolers' give-a-number estimates as coming from a logarithmic to linear shift in representations of numerical magnitude (Opfer et al., 2010). Consistent with this idea, the correlation between accuracy and linearity measures was considerable [$r(36) = 0.89$, $p < 0.001$; **Table 2**], even when age was controlled [$r(35) = 0.88$, $p < 0.001$]. We also observed the median number of chips given by 3-year-olds increased logarithmically with the number of chips requested ($\lambda = 0.68$), whereas chips given by 4- and 5-year-olds increased linearly with the number requested (4s: $\lambda = 0.19$, 5s: 0.02). Thus, at the group level, all three age groups appeared to map numeric magnitudes at least approximately to the number of chips they provided, with superiority of the linear to the logarithmic functions increasing with age.

We next tested for an approximate mapping between number and magnitude on the give-a-number test at the individual level. As done on the number-line estimation test, we regressed the number of chips requested by an experimenter against the number of chips given by each child; the proportion of non-mapping children, whose estimates were not explainable by a MLLM, was calculated. Children were categorized into either non-mappers or mappers using the goodness of fit criterion

parameters of sub-models in MCPMs (i.e., w_1 , w_2 , and w_3) that are thought to capture developmental changes presented no systematic patterns that might reflect age-related improvement.

TABLE 1 | Mean performance (and SDs) by tasks for each age group.

Task	Measure	Age			
		3-year-old	4-year-old	5-year-old	Total
Number-line	MPAE	32.19 (11.75)	23.68 (13.59)	12.76 (6.56)	23.16 (13.47)
	Log component	0.91 (0.28)	0.61 (0.44)	0.29 (0.36)	0.61 (0.44)
Give-a-number	MPAE	20.38 (12.26)	13.35 (13.19)	4.32 (5.99)	12.92 (12.64)
	Log component	0.53 (0.47)	0.40 (0.43)	0.13 (0.28)	0.36 (0.42)
Counting	MPAE	14.62 (10.92)	8.13 (8.54)	3.85 (4.39)	9.00 (9.38)
	Log component	0.47 (0.47)	0.09 (0.12)	0.03 (0.09)	0.20 (0.35)
Color recall	Accuracy (%)	79.81 (25.79)	86.54 (13.94)	96.88 (5.65)	87.5 (18.38)
Number recall	Accuracy (%)	48.08 (26.44)	55.77 (23.73)	76.04 (24.11)	59.54 (26.86)
	MPAE	10.78 (7.26)	10.94 (10.19)	4.48 (6.39)	8.84 (8.48)

MPAE, mean percent absolute error.

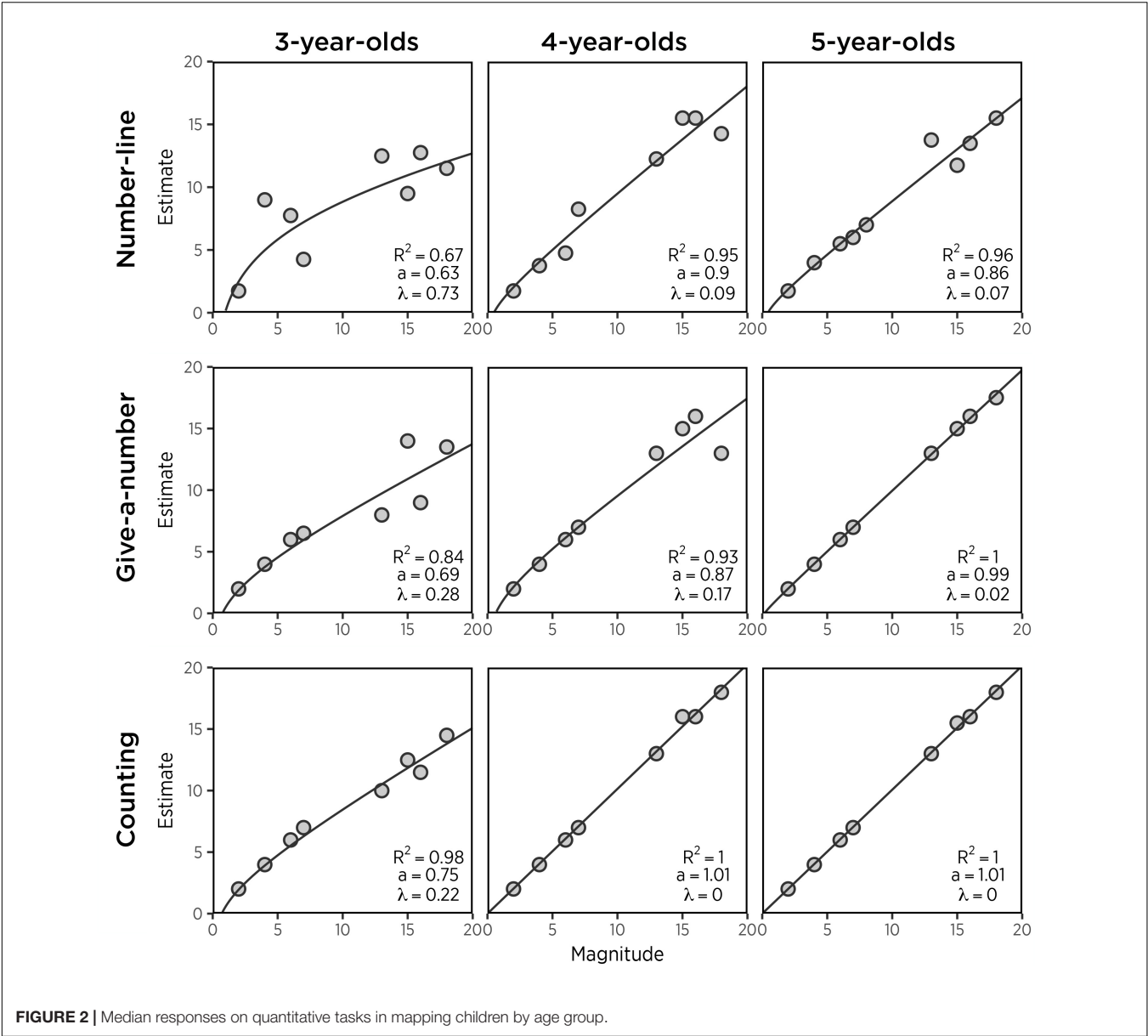


TABLE 2 | Correlations among numeric memory measures and predictor variables.

	1	2	3	4	5	6	7	8	9
1. Number recall accuracy (%)									
2. Number recall MPAE	−0.75***								
3. Age	0.48**	−0.32*							
4. Color recall accuracy (%)	0.21	−0.29	0.47**						
5. NL MPAE	−0.48**	0.42**	−0.53**	−0.23					
6. GAN MPAE	−0.44**	0.43**	−0.59***	−0.57***	0.68***				
7. Counting MPAE	−0.13	0.35*	−0.47**	−0.53**	0.49**	0.73***			
8. NL log component	−0.46**	0.42**	−0.54**	−0.19	0.64***	0.41*	0.31		
9. GAN log component	−0.38*	0.41*	−0.42**	−0.55***	0.55***	0.89***	0.66***	0.29	
10. Counting log component	−0.18	0.11	−0.52**	−0.55***	0.53**	0.76***	0.76***	0.19	0.71***

MPAE, mean percent absolute error; NL, number-line task; GAN, give-a-number task. Asterisks indicate levels of significance (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

(i.e., R^2). In total, only 10.3% of all preschoolers showed non-significant correlation between the chips given and the number requested (three 3-year-olds and one 4-year-old). These results indicate a better understanding of the task than observed for number-line estimation, though many younger children still failed to map symbolic number to the number of chips even in an approximate fashion.

To test the hypothesized learning sequence, we next evaluated age as a predictor of both linearity and accuracy measures. Mapping children ($n = 35$) showed a marginally significant improvement in linearity, $b = -0.17$, $t(33) = -2.02$, $p = 0.052$, and accuracy, $b = -7.08$, $t(33) = -3.71$, $p < 0.001$. Age explained 11% of variance in linearity, $F(1,33) = 4.07$, $p = 0.052$, and 29.4% of variance in accuracy, $F(1,33) = 13.77$, $p < 0.001$. Median estimates increased more logarithmically with actual value in 3- and 4-year-old mappers (3s: $\lambda = 0.28$, 4s: $\lambda = 0.17$), compared to 5-year-old mapping children whose median estimates were almost perfectly linear ($\lambda = 0.02$) (Figure 2). Thus, like the results from number-line estimation, results suggest that preschoolers who approximately map number to magnitude show a log-to-linear progression in the representation used.

Counting Task

Counting accuracy improved significantly with age, $b = -5.83$, $t(36) = -3.20$, $p < 0.01$, as did linearity, $b = -0.24$, $t(36) = -3.63$, $p < 0.001$ (Figure 1A). Age also explained a significant percentage of variance in accuracy, 22.1%, $F(1,36) = 10.22$, $p < 0.01$, and in linearity, 26.8%, $F(1,36) = 13.25$, $p < 0.001$. The average MPAE for all children was 9.00 ($SD = 9.38$), and average logarithmic component was 0.20 ($SD = 0.35$) (Table 1). The MPAE was 14.62 ($SD = 10.92$) for 3-year-olds, 8.13 ($SD = 8.54$) for 4-year-olds, and 3.85 ($SD = 4.39$) for 5-year-olds. The average logarithmic component for 3-year-olds was 0.47 ($SD = 0.47$), 0.09 ($SD = 0.12$) for 4-year-olds, and 0.03 ($SD = 0.09$) for 5-year-olds.

Previous work had explained age-related changes in accuracy of preschoolers' counting as coming from procedural knowledge rather than representational change (Briars and Siegler, 1984; Wynn, 1990). Somewhat surprisingly, however, we found significant relations between accuracy and linearity in counting no matter whether age effects were partialled out [$r(36) = 0.76$, $p < 0.001$ for the zero-order correlation; $r(35) = 0.69$, $p < 0.001$ for the partial correlation]. In addition, median counts of 3-year-olds increased more logarithmically with the number of chips presented ($\lambda = 0.38$), whereas median counts of 4- and 5-year-olds increased perfectly linearly with chips presented (4s: $\lambda = 0$, 5s: $\lambda = 0$). Thus, at the group level, all three age groups appeared to map numeric magnitudes at least approximately to the number of chips they provided, with superiority of the linear to the logarithmic functions increasing with age.

Because, as we have seen, that analyses of group data are not always consistent with analyses of individual performance, we next tested for an approximate mapping between number and magnitude at the individual level, as we did on the two estimation tests. As in number-line and give-a-number tasks, children were divided into two groups, mapping and non-mapping groups, based on fitting of a MLLM to their responses. Only one 3-year-old child was found to have no relation

between the chips given and the number requested. Without the one non-mapper, the age effects in linearity and accuracy still stayed significant at both individual and group levels. For example, after taking out the non-mapper, median responses of 3-year-olds still increased more logarithmically with actual given magnitudes ($\lambda = 0.22$), whereas 4- and 5-year-olds' median responses were completely linear (4s: $\lambda = 0$, 5s: $\lambda = 0$), suggesting log-to-linear developmental shifts in counting (Figure 2). Taken together, these results show that in general most 3- to 5-year-old preschoolers were capable of approximately mapping the number of chips to their counts and that there were developmental progresses in their mapping.

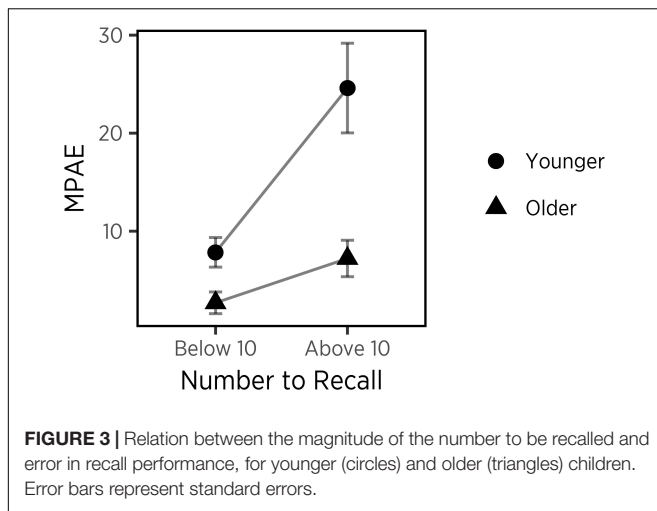
Is the developmental path of counting different from that of number-line and give-a-number estimation that requires deep understanding of number-to-magnitude mappings? Our results from accuracy and linearity in counting above suggest that even though number-to-magnitude mapping is not necessary in counting, it follows log-to-linear shifts as in other estimation tasks (Figure 1). However, counting appeared to be more accurate and linear than the other tasks. As shown in Figure 2, median responses of mapping children increased more linearly from the age three and reached perfect linearity earlier in counting than number-line and give-a-number estimation. Consistent with median responses, individual children performed better in counting than the other tasks (MPAE: $M_{NL} = 23.16$, $M_{GAN} = 12.92$, $M_{count} = 9.0$; λ : $M_{NL} = 0.61$, $M_{GAN} = 0.40$, $M_{count} = 0.20$). Whereas counting accuracy correlated with accuracy of estimation tasks, $r_s = 0.49$ – 0.73 , $p < 0.01$, linearity in counting showed no association with number-line estimation, but only with give-a-number estimation, $r(36) = 0.71$, $p < 0.001$ (Table 2). Together, even if counting shares the developmental trajectory with estimation, counting improves faster than estimation.

Number/Color Recall Task

We next examined age-related improvements in the percentage of colors recalled (Figure 1B). Color recall also showed a strong effect of age, $b = 11.37$, $t(36) = 3.58$, $p < 0.01$, with age explaining 22.9% of variance, $F(1,36) = 10.09$, $p < 0.01$. Overall, the percentage of colors recalled accurately improved with age, (3s, $M = 79.81\%$, $SD = 25.79\%$; 4s, $M = 86.54\%$, $SD = 13.94\%$; 5s, $M = 96.88\%$, $SD = 5.65\%$) (Table 1). To compare color vs. number memories, we computed the percentage of numbers correctly recalled (e.g., correct if children recall 5 and incorrect if they recalled 6 after counting 5). When two types of recall accuracy were compared, color recall was superior to number recall, $t(37) = 5.89$, $p < 0.001$, presumably due to its greater temporal proximity.

To calculate MPAE for the number recall task, we took the average of the Percentage Absolute Error (PAE), or $(|\text{Number Counted} - \text{Number Remembered}|/20) \times 100$, across all trials for children. Thus, if a child (correctly or incorrectly) said there were 12 chips on a card and then recalled there being 13 chips, PAE would be 5%.

To examine development of numeric recall, we carried out a regression between age and the percentage of numbers recalled perfectly, as well as between age and MPAE in recalling the



numbers they initially counted. As expected, the recall of the exact numbers improved with age, $b = 17.07$, $t(36) = 3.29$, $p < 0.01$, and age explained 23.1% of variance, $F(1,36) = 10.8$, $p < 0.01$. The mean percentage of exact number recall trials was 48.08% ($SD = 26.44\%$) for 3-year-olds, 55.77% ($SD = 23.73\%$) for 4-year-olds, and 76.04% ($SD = 24.11\%$) for 5-year-olds (Table 1). More importantly, younger children recalled numbers that were more deviated from correct numbers, whereas older children retrieved numbers more accurately (Figure 1B). Recall MPAE improved with age, $b = -3.63$, $t(36) = -2.05$, $p < 0.05$, and age explained 10.1% of variance, $F(1,36) = 4.21$, $p < 0.05$. The average MPAE for all children was 8.84 ($SD = 8.48$). The mean of individuals' MPAE was 10.78 ($SD = 7.26$) for 3-year-olds, 10.94 ($SD = 10.19$) for 4-year-olds, and 4.48 ($SD = 6.39$) for 5-year-olds. As expected, then, we observed age-related increases in the percentage of numbers recalled perfectly and decreases in MPAE of numeric recall.

To test the representational change account of recall more directly, we examined memory MPAE for an interaction of age and numeric magnitude. This interaction is predicted uniquely by the idea that the subjective magnitudes of numbers change with age. According to the representational change account, children should show good recall for small numbers regardless of representation. However, for large numbers, only children who possess a linear representation would be able to distinguish them from one another, whereas children with a logarithmic representation would show more erroneous memory for large numbers due to greater overlaps in representation. To test this, we divided number recalls into two categories based on the number of digits that a stimulus contained (i.e., numbers below 10 vs. numbers above 9) to control for visual features shared by single-digit and two-digit numbers. Children were also divided into two groups relative to the median split of ages for children in the study (4.43 years old). As predicted, a mixed ANOVA showed a main effect of numeric magnitude on recall accuracy, $F(1,34) = 17.79$, $p < 0.001$, a main effect of age group on recall accuracy, $F(1,34) = 18.90$, $p < 0.001$. More importantly, an interaction effect of numeric magnitude and age group on

recall accuracy, $F(1,34) = 5.59$, $p < 0.05$ (see Figure 3). *Post hoc* comparisons using Bonferroni's adjustment revealed that younger children generated significantly greater errors for larger numbers ($M = 24.60$, $SD = 3.45$) than for smaller numbers ($M = 8.54$, $SD = 1.39$), $p < 0.001$, whereas errors of smaller numbers ($M = 2.69$, $SD = 1.31$) did not differ from those of larger numbers ($M = 7.22$, $SD = 3.26$) in older children, $p = 0.19$. Taken together, the results suggest that younger children with a logarithmic representation produced greater errors for larger numbers that have more overlaps with other numbers in representation, supporting the representational change account.

Logarithmic Compression in Numerical Tasks as Predictors of Memory Performance

Might improvement in memory accuracy—like improvement in accuracy and linearity of numerical estimates, give-a-number estimates, and counting—be related to increasingly linear representations of numerical value? Correlations among tasks show that it might be the case. As shown in Table 2, numeric memory measured with percent correct and MPAE was strongly associated with accuracy and linearity of number-line and give-a-number estimates, but weakly correlated with those of counting. To test this more closely, we conducted multiple regression analyses on the mean percent deviation between recalled and correct numbers (MPAE) in the recall task. In the analyses, children's age, percent correct responses in color recall, and average value of numbers to recall were entered in a regression model to control for the influences of the variables. The mean of to-be-recalled numbers varied across children because the stimuli were generated by individual children in the counting task. In addition to the three predictors, MPAs from three numerical tasks were included in a regression model in order to examine the unique contributions of the numeric tasks to the MPAE in number recall simultaneously.

The model accounted for a significant amount of variance in recall MPAE (64%), $F(6,31) = 9.14$, $p < 0.001$. The errors in children's number recall were explained by the mean values of numbers that children produced themselves to recall later [$b = 3.02$, $\beta = 0.77$, $t(31) = 5.97$, $p < 0.001$]. Children who produced larger magnitudes on average in counting tended to produce more erroneous responses in number recall. Interestingly, the accuracy of number-line estimation was another significant predictor for the number recall accuracy [$b = 0.22$, $\beta = 0.35$, $t(31) = 2.20$, $p < 0.04$]. Whereas give-a-number MPAE was marginally predictive of recall performance [$b = 0.30$, $\beta = 0.41$, $t(31) = 1.95$, $p = 0.06$], MPAE in counting did not explain errors in number recall [$b = -0.04$, $\beta = -0.04$, $t(31) = -0.27$, $p = 0.79$]. Interestingly, whereas age was a significant predictor for number memory in a simple regression ($b = -3.63$, $p < 0.05$), the age effects were not evident in the multiple regression, where the age variable was tested with five other predictors [$b = -1.79$, $\beta = -0.16$, $t(31) = -1.11$, $p = 0.28$]. Neither was the color recall accuracy (percent correct) a significant predictor for the number recall performance [$b = -8.65$, $\beta = -0.19$, $t(31) = -1.31$, $p = 0.20$].

Using the same analyses, we next tested whether linearity of some tasks better predicted numerical memory than that of other tasks. Individuals' MPAEs from the three numeric tasks were replaced with their respective logarithmic components (λ) as predictors. More than 66% of variance in number recall accuracy was addressed by the six predictors (age, percent of correct responses in color recall, average value of numbers to recall, and λ s for number-line estimation, give-a-number estimation, and counting), $F(6,31) = 10.44$, $p < 0.001$. Again, age and accuracy in color memory did not predict number recall significantly [$b = -1.82$, $\beta = -0.16$, $t(31) = -1.11$, $p = 0.28$ for age, $b = -6.39$, $\beta = -0.14$, $t(31) = -1.05$, $p = 0.30$ for color memory]. On the other hand, the mean number that children produced to remember has a considerable contribution to MPAE in number recall [$b = 2.98$, $\beta = 0.76$, $t(31) = 5.38$, $p < 0.001$]. More importantly, degrees of linearity in both number-line and give-a-number estimation accounted for performance in number recall [$b = 7.48$, $\beta = 0.39$, $t(31) = 3.02$, $p < 0.01$ for number-line linearity, $b = 9.00$, $\beta = 0.44$, $t(31) = 2.82$, $p < 0.01$ for give-a-number linearity]. The logarithmic component in counting did not associate with number recall in a significant way [$b = 1.25$, $\beta = 0.05$, $t(31) = 0.27$, $p = 0.79$].

Next, extending the multiple regressions, we conducted mixed-effects modeling on trial-to-trial PAEs of number recall with varying intercepts for participants to investigate relations between number memory and numerical tasks. The mixed-effects model allows for examining average (fixed) effects of numerical tasks on number memory across children while also accounting for individual differences among children. In the analysis, fixed effects included children's age, color recall accuracy, number to recall, and MPAEs from number-line estimation, give-a-number estimation, and counting. The p -values for fixed effects were computed using Satterthwaite's approximation method (Satterthwaite, 1941) to define denominator degrees of freedom for the t -test. Intercepts were treated as random at the participant level to control for inter-individual variability. When the effects of the six variables were averaged over all children, the number that children produced to recall was the only significant fixed effect [$b = 1.52$, $\beta = 0.52$, $t(31) = 10.58$, $p < 0.001$], implying that PAE in number recall increased with the magnitude of numbers to recall. Neither age nor accuracy measures of three numerical tasks showed significant effects on accuracy in number recall.

Another linear mixed-effects analysis was conducted on PAEs for every trial in number recall. The model was identical to the one described above except that the MPAEs from the numeric tasks were replaced with their respective logarithmic components (λ). Again, fixed effects of numbers to recall was significant [$b = 1.51$, $\beta = 0.52$, $t(31) = 10.47$, $p < 0.001$], indicating that number recall accuracy varied depending on the magnitude of to-be-recalled numbers. Using the linearity instead of accuracy, the model showed significant effects of logarithmic components of number-line estimation and give-a-number tasks [$b = 5.94$, $\beta = 0.14$, $t(31) = 2.26$, $p < 0.05$ for number-line estimation, $b = 10.59$, $\beta = 0.23$, $t(31) = 3.13$, $p < 0.01$ for give-a-number]. The results suggest that the more logarithmic in the two numerical tasks, the fuzzier number recall performance, and that the linearity indices reliably predict number memory. The linearity

measure from counting was not significant [$b = -4.54$, $\beta = -0.08$, $t(31) = -1.04$, $p = 0.30$]. Again, age and accuracy of color recall did not contribute to number recall.

DISCUSSION

Previous work has indicated that development of linear representations of numerical magnitudes profoundly expands children's quantitative thinking (Opfer and Siegler, 2012). It improves children's ability to estimate the positions of numbers on number lines (Siegler and Opfer, 2003; Siegler and Booth, 2004; Opfer and Siegler, 2007), to estimate the measurements of continuous quantities (Thompson and Siegler, 2010) and the quantity of discrete objects (Opfer et al., 2010), to categorize numbers according to size (Laski and Siegler, 2007; Opfer and Thompson, 2008), and to estimate and learn the answers to arithmetic problems (Booth and Siegler, 2008; Kim and Opfer, 2017; Qin et al., 2017). Recent work has also indicated that the logarithmic-to-linear shift is associated with improved memory for numbers (Thompson and Siegler, 2010; Thompson and Opfer, 2016).

In this paper, we took a critical look at the representational change theory of development of numerical recall. We were particularly interested in whether it could account for changes in ability to recall single numbers that children themselves produced. This issue is important because previous work could not rule out the influence of memory span on numerical memory. Further, by examining numbers that children themselves produced, we could eliminate the possibility that preschoolers with non-linear numerical-magnitude representations were simply poor at remembering unfamiliar numbers. Thus, we sought to provide a robust test of the theory.

Consistent with the representational change account, we found preschoolers' recall of a single number to be closely tied to the linearity of their mapping between numeric symbols and quantities. Indeed, this connection to preschoolers' numerical recall was even beyond what would be expected based solely on their age or memory for other items, such as self-generated color words. Further, consistent with the hypothesis that young children are unable to correctly recall large numbers due to increasing semantic similarity among large numeric magnitudes, we found that young preschoolers' memory for small numbers was nearly equivalent to that of older preschoolers' memory, whereas older preschoolers recalled large numbers much more accurately than younger preschoolers. An intriguing question for future research is the extent to which the semantic similarity of numbers co-varies with other forms of similarity (e.g., phonological or visual form; Cohen, 2009) and which type of similarity best predicts numeric recall.

In addition, our results showed that children's performance in number recall was predicted by accuracy and, more reliably, by linearity in number-to-quantity mapping tasks—i.e., number-line estimation and give-a-number tasks—as well as by the magnitude of numbers to recall. Surprisingly, the effects of age and memory for non-numerical items on number recall were not evident when all predictors were considered simultaneously. The findings remained consistent no matter whether the effects were

examined with single-level or multi-level analyses. Together, the findings provide strong evidence for the representational change account.

Theoretically, the effect of numerical magnitude on recall accuracy is important because it suggests a new way to integrate findings regarding development of memory and numerical cognition. That is, both areas of research strongly suggest that long-duration representations of numerical magnitude are “fuzzy” and approximate (Gallistel and Gelman, 1992; Brainerd and Reyna, 1993; Dehaene and Changeux, 1993; Brainerd and Gordon, 1994). However, unlike the findings integrated by the FTT, findings on development of numerical estimation suggest that the internal magnitudes associated with numbers change over development from a logarithmic to a linear association (Siegler and Opfer, 2003; Opfer and Siegler, 2007; Siegler, 2016), with the result that the distinctiveness of the representation of numeric magnitudes is initially larger for small numbers than large ones. The implication of this view for numeric recall comes from the general finding that the probability of recall is positively related to the distinctiveness of the representation in memory (Greene and Crowder, 1984), with the apparently correct prediction that recall accuracy would be initially greater for small numbers than large numbers and that this difference would decline with age. Previous work has demonstrated that adults produce non-linear estimates of very large numbers (e.g., a million) as young children do for small numbers (Landy et al., 2013, 2017). Given that compressive number-to-magnitude mapping is not limited to children, whether adults’ memories for numbers are also subject to magnitudes is an interesting question for future research.

Beyond demonstrating that linear numerical-magnitude representations are associated with improved memory for numbers, the present results also help to explain the positive relation between linear numeric magnitude representations and arithmetic proficiency (Booth and Siegler, 2008;

Kim and Opfer, 2017; Qin et al., 2017). That is, if developing linear numerical-magnitude representations improves memory for single numbers (e.g., four chips) as well as multiple numbers presented in vignettes (e.g., four cows, six cows), it is highly likely that it also improves memory for numbers in other contexts, such as memorizing arithmetic facts (e.g., 4 cows + 6 cows = 10 cows). In this way, the present results suggest a plausible explanation for the observed association between numerical estimation and mathematics course grades (Booth and Siegler, 2008; Halberda et al., 2008), and it suggests that numerical memory may moderate this link. Although this account is admittedly speculative, we believe it is an important issue for future research.

AUTHOR CONTRIBUTIONS

JO designed the experiments, analyzed the data, and co-wrote the manuscript. DK and CY analyzed data and contributed to writing of the manuscript. FM collected the data.

FUNDING

This research was supported by the Institute for Educational Sciences (United States Department of Education), R305A160295.

ACKNOWLEDGMENTS

The authors would like to thank Clarissa Thompson for comments on earlier draft of this paper. A portion of these data was presented at the 33rd Annual Meeting of the Cognitive Science Society, Boston, 20–23rd of July, 2011.

REFERENCES

- Anobile, G., Cicchini, G. M., and Burr, D. C. (2012). Linear mapping of numbers onto space requires attention. *Cognition* 122, 454–459. doi: 10.1016/j.cognition.2011.11.006
- Barth, H. C., and Paladino, A. M. (2011). The development of numerical estimation: evidence against a representational shift. *Dev. Sci.* 14, 125–135. doi: 10.1111/j.1467-7687.2010.00962.x
- Berteletti, I., Lucangeli, D., Piazza, M., Dehaene, S., and Zorzi, M. (2010). Numerical estimation in preschoolers. *Dev. Psychol.* 46, 545–551. doi: 10.1037/a0017887
- Booth, J. L., and Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Dev.* 79, 1016–1031. doi: 10.1111/j.1467-8624.2008.01173.x
- Brainerd, C. J., and Gordon, L. L. (1994). Development of verbatim and gist memory for numbers. *Dev. Psychol.* 30, 163–177. doi: 10.1037/0012-1649.30.2.163
- Brainerd, C. J., and Reyna, V. F. (1993). Memory independence and memory interference in cognitive development. *Psychol. Rev.* 100, 42–67. doi: 10.1037/0033-295X.100.1.42
- Briars, D. J., and Siegler, R. S. (1984). A featural analysis of preschoolers’ counting knowledge. *Dev. Psychol.* 20, 607–618. doi: 10.1037/0012-1649.20.4.607
- Cantlon, J. F., Cordes, S., Libertus, M. E., and Brannon, E. M. (2009). Comment on “Log or Linear? Distinct intuitions of the number scale in western and amazonian indigene cultures”. *Science* 323, 38; author reply 38. doi: 10.1126/science.1164773
- Cohen, D. J. (2009). Integers do not automatically activate their quantity representation. *Psychon. Bull. Rev.* 16, 332–336. doi: 10.3758/PBR.16.2.332
- Cohen, D. J., and Quinlan, P. T. (2018). The log–linear response function of the bounded number-line task is unrelated to the psychological representation of quantity. *Psychon. Bull. Rev.* 25, 447–454. doi: 10.3758/s13423-017-1290-z
- Cohen, D. J., and Sarnecka, B. W. (2014). Children’s number-line estimation shows development of measurement skills (not number representations). *Dev. Psychol.* 50, 1640–1652. doi: 10.1037/a0035901
- Dehaene, S., and Changeux, J.-P. (1993). Development of elementary numerical abilities: a neuronal model. *J. Cogn. Neurosci.* 5, 390–407. doi: 10.1162/jocn.1993.5.4.390
- Dempster, F. N. (1981). Memory span: sources of individual and developmental differences. *Psychol. Bull.* 89, 63–100. doi: 10.1037/0033-2909.89.1.63
- Ebersbach, M., Luwel, K., Frick, A., Onghena, P., and Verschaffel, L. (2008). The relationship between the shape of the mental number line and familiarity with numbers in 5- to 9-year-old children: evidence for a segmented linear model. *J. Exp. Child Psychol.* 99, 1–17. doi: 10.1016/j.jecp.2007.08.006
- Gallistel, C. R., and Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition* 44, 43–74. doi: 10.1016/0010-0277(92)90050-R
- Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent, L., and Numtee, C. (2007). Cognitive mechanisms underlying achievement deficits in children with

- mathematical learning disability. *Child Dev.* 78, 1343–1359. doi: 10.1111/j.1467-8624.2007.01069.x
- Greene, R. L., and Crowder, R. G. (1984). Effects of semantic similarity on long-term recency. *Am. J. Psychol.* 3, 441–449. doi: 10.2307/1422530
- Halberda, J., Mazocco, M., and Feigenson, L. (2008). Individual differences in nonverbal number acuity predict maths achievement. *Nature* 455, 665–668. doi: 10.1038/nature07246
- Kim, D., and Opfer, J. E. (2017). A unified framework for bounded and unbounded numerical estimation. *Dev. Psychol.* 53, 1088–1097. doi: 10.1037/dev0000305
- Landy, D., Charlesworth, A., and Ottmar, E. (2017). Categories of large numbers in line estimation. *Cogn. Sci.* 41, 326–353. doi: 10.1111/cogs.12342
- Landy, D., Silbert, N., and Goldin, A. (2013). Estimating large numbers. *Cogn. Sci.* 37, 775–799. doi: 10.1111/cogs.12028
- Laski, E. V., and Siegler, R. S. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child Dev.* 76, 1723–1743. doi: 10.1111/j.1467-8624.2007.01087.x
- Le Corre, M., Van de Walle, G., Brannon, E. M., and Carey, S. (2006). Revisiting the competence/performance debate in the acquisition of the counting principles. *Cogn. Psychol.* 52, 130–169. doi: 10.1016/j.cogpsych.2005.07.002
- Lee, M. D., and Sarnecka, B. W. (2010). A model of knower-level behavior in number-concept development. *Cogn. Sci.* 34, 51–67. doi: 10.1111/j.1551-6709.2009.01063.x
- Moeller, K., Pixner, S., Kaufmann, L., and Nuerk, H.-C. (2009). Children's early mental number line: logarithmic or decomposed linear? *J. Exp. Child Psychol.* 103, 503–515. doi: 10.1016/j.jecp.2009.02.006
- Opfer, J. E., and Siegler, R. S. (2007). Representational change and children's numerical estimation. *Cogn. Psychol.* 55, 169–195. doi: 10.1016/j.cogpsych.2006.09.002
- Opfer, J. E., and Siegler, R. S. (2012). "Development of quantitative thinking," in *Oxford Handbook of Thinking and Reasoning*, eds K. Holyoak and R. Morrison (New York, NY: Oxford University Press).
- Opfer, J. E., Siegler, R. S., and Young, C. J. (2011). The powers of noise-fitting: reply to Barth and Paladino. *Dev. Sci.* 14, 1194–1204. doi: 10.1111/j.1467-7687.2011.01070.x
- Opfer, J. E., and Thompson, C. A. (2008). The trouble with transfer: insights from microgenetic changes in representations of numerical magnitude. *Child Dev.* 79, 788–804. doi: 10.1111/j.1467-8624.2008.01158.x
- Opfer, J. E., Thompson, C. A., and Furlong, E. (2010). Early development of spatial-numeric associations: evidence from spatial and quantitative performance of preschoolers. *Dev. Sci.* 13, 761–771. doi: 10.1111/j.1467-7687.2009.00934.x
- Opfer, J. E., Thompson, C. A., and Kim, D. (2016). Free versus anchored numerical estimation: a unified approach. *Cognition* 149, 11–17. doi: 10.1016/j.cognition.2015.11.015
- Qin, J., Kim, D., and Opfer, J. E. (2017). "Varieties of numerical estimation: a unified framework," in *Proceedings of the 39th Annual Meeting of the Cognitive Science Society*, eds G. Gunzelmann, A. Howes, T. Tenbrink, and E. J. Davelaar (Austin, TX: Cognitive Science Society), 2943–2948.
- Reyna, V. F., Nelson, W. L., Han, P. K., and Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychol. Bull.* 135, 943–973. doi: 10.1037/a0017327
- Sarnecka, B. W., and Carey, S. (2008). How counting represents number: what children must learn and when they learn it. *Cognition* 108, 662–674. doi: 10.1016/j.cognition.2008.05.007
- Satterthwaite, F. E. (1941). Synthesis of variance. *Psychometrika* 6, 309–316. doi: 10.1007/BF02288586
- Schacter, D. L., Norman, K. A., and Koustaal, W. (1998). The cognitive neuroscience of constructive memory. *Annu. Rev. Psychol.* 49, 289–318. doi: 10.1146/annurev.psych.49.1.289
- Sella, F., Berteletti, I., Lucangeli, D., and Zorzi, M. (2017). Preschool children use space, rather than counting, to infer the numerical magnitude of digits: evidence for a spatial mapping principle. *Cognition* 158, 56–67. doi: 10.1016/j.cognition.2016.10.010
- Siegler, R. S. (2016). Magnitude knowledge: the common core of numerical development. *Dev. Sci.* 19, 341–361. doi: 10.1111/desc.12395
- Siegler, R. S., and Booth, J. L. (2004). Development of numerical estimation in young children. *Child Dev.* 75, 428–444. doi: 10.1111/j.1467-8624.2004.00684.x
- Siegler, R. S., and Opfer, J. E. (2003). The development of numerical estimation: evidence for multiple representations of numerical quantity. *Psychol. Sci.* 14, 237–243. doi: 10.1111/1467-9280.02438
- Siegler, R. S., and Ramani, G. B. (2009). Playing linear number board games—but not circular ones—improves low-income preschoolers' numerical understanding. *J. Educ. Psychol.* 101:545. doi: 10.1037/a0014239
- Siegler, R. S., Thompson, C. A., and Opfer, J. E. (2009). The logarithmic-to-linear shift: one learning sequence, many tasks, many ages. *Mind Brain Educ.* 3, 143–150. doi: 10.1111/j.1751-228X.2009.01064.x
- Slusser, E. B., Santiago, R. T., and Barth, H. C. (2013). Developmental change in numerical estimation. *J. Exp. Psychol. Gen.* 142, 193–208. doi: 10.1037/a0028560
- Thompson, C. A., and Opfer, J. E. (2008). Costs and benefits of representational change: effect of context on age and sex differences in magnitude estimation. *J. Exp. Child Psychol.* 101, 20–51. doi: 10.1016/j.jecp.2008.02.003
- Thompson, C. A., and Opfer, J. E. (2010). How 15 hundred is like 15 cherries: effect of progressive alignment on representational changes in numerical cognition. *Child Dev.* 81, 1768–1786. doi: 10.1111/j.1467-8624.2010.01509.x
- Thompson, C. A., and Opfer, J. E. (2016). Learning linear spatial-numeric associations improves accuracy of memory for numbers. *Front. Psychol.* 7:24. doi: 10.3389/fpsyg.2016.00024
- Thompson, C. A., and Siegler, R. S. (2010). Linear numerical-magnitude representations aid children's memory for numbers. *Psychol. Sci.* 21, 1274–1281. doi: 10.1177/0956797610378309
- Wagner, J. B., and Johnson, S. C. (2011). An association between understanding cardinality and analog magnitude representations in preschoolers. *Cognition* 119, 10–22. doi: 10.1016/j.cognition.2010.11.014
- Wynn, K. (1990). Children's understanding of counting. *Cognition* 36, 155–193. doi: 10.1016/0010-0277(90)90003-3
- Young, C. J., and Opfer, J. E. (2011). Psychophysics of numerical representation: a unified approach to single- and multi-digit magnitude estimation. *J. Psychol.* 219, 58–63. doi: 10.1027/2151-2604/a000047

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Opfer, Kim, Young and Marciani. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



English and Chinese Children's Performance on Numerical Tasks

Ann Dowker^{1*} and Anthony M. Li²

¹ Department of Experimental Psychology, Oxford University, Oxford, United Kingdom, ² Somerville College, Oxford University, Oxford, United Kingdom

East Asian pupils have consistently outperformed Western pupils in international comparisons of mathematical performance at both primary and secondary school level. It has sometimes been suggested that a contributory factor is the transparent counting systems of East Asian languages, which may facilitate number representation. The present study compared 35 7-year-old second-year primary school children in Oxford, England and 40 children of similar age in Hong Kong, China on a standardized arithmetic test; on a two-digit number comparison test, including easy, misleading and reversible comparisons; and on a number line task, involving placing numbers in the appropriate position on four number lines: 1–10, 1–20, 1–100, and 1–1000. The Chinese children performed significantly better than the English children on the standardized arithmetic test. They were faster but not significantly more accurate on the Number Comparison and Number Line tasks. There were no interactions between language group and comparison type on the number comparison task, though the performance of both groups was faster on easy pairs than those where there was conflict between the relative magnitudes of the tens and the units. Similarly, there were no interactions between group and number line range, though the performance of both groups was influenced by the range of the number line. The study supports the view that counting systems affect aspects of numerical abilities, but cannot be the full explanation for international differences in mathematics performance.

OPEN ACCESS

Edited by:

Krzysztof Cipora,
University of Tübingen, Germany

Reviewed by:

Andrea Bender,
University of Bergen, Norway
Amandine Van Rinsveld,
Free University of Brussels, Belgium

*Correspondence:

Ann Dowker
ann.dowker@psy.ox.ac.uk

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 06 April 2018

Accepted: 18 December 2018

Published: 05 February 2019

Citation:

Dowker A and Li AM (2019)
English and Chinese Children's
Performance on Numerical Tasks.
Front. Psychol. 9:2731.
doi: 10.3389/fpsyg.2018.02731

Keywords: primary school children, mathematical development, number line tasks, number comparison, cross-cultural research, counting system

INTRODUCTION

Recent large-scale cross-national comparisons of mathematical abilities (Askew et al., 2010; Sturman, 2015; Mullis et al., 2016a,b) have shown that East Asian countries like China, Japan, South Korea, and Singapore are usually at the top of international comparisons of mathematics performance. Most studies have found an East Asian advantage in mathematical performance in multiple age groups, starting from preschool (Miller et al., 1995; Geary et al., 1996).

There are many possible reasons for East Asian children's particularly high performance on these tasks. These include differences in teaching methods: indeed, in recent years, United Kingdom schools have been seeking to develop and use materials and approaches similar to those used in Shanghai and Singapore. Different researchers and policymakers emphasize different aspects of the teaching approaches that they see as beneficial. Some emphasize greater subject knowledge and understanding by East Asian teachers, reinforced by extensive continuous professional

development (Ma, 1999); some emphasize greater attempts to foster conceptual understanding (Perry et al., 1993; Stigler et al., 1996; Ma, 1999) some emphasize greater focus on rote learning (Gibb, 2014); some emphasize the ‘mastery’ approach whereby fewer areas within mathematics are covered, but in greater depth, and teachers endeavor to ensure that all pupils in a class have understood one topic before moving on to the next (Jerrim and Vignoles, 2015). Additionally, East Asian pupils often devote more hours per day to mathematics (and some other academic subjects) in school and in homework than those in many other countries. Also, East Asians may value mathematics more than those in many other countries; and appear to place more value on academic achievement in general, and to attribute success more to effort, than many Westerners (Hess et al., 1987; Stigler et al., 1996; Wong et al., 2001).

One further explanation that has been proposed for East Asian children’s relatively high performance in mathematics is that their languages have highly transparent counting systems (Miura et al., 1988). In Chinese, for instance, counting from ten onwards takes the form of A-ten-B, and then A-hundred-B-ten-C. Twelve in Chinese is 十二 (shi-er), which translates as ten-two; Sixty-two in Chinese is 六十二 (liu-shi-er), which translates as six-ten-two.

By contrast, the English counting system is more opaque. There are three major irregularities in the English counting system below 100. Firstly, the numbers ‘eleven’ and ‘twelve’ give children no suggestion of the cardinality of the number. In contrast with the Chinese counting words, there are no indications of number values within these English words – that eleven means ten plus one, while twelve means ten plus two. Secondly, the teen words are inverted in relation to with Arabic numerals; e.g., ‘sixteen’ is inverted compared to the Arabic ‘16’ and the Chinese +7 (shi-liu, literally ten-six); the same applies to numbers from thirteen to nineteen. Thirdly, the teen words sound similar to the numbers that are multiples of ten, e.g., ‘sixteen’ sounds like ‘sixty,’ which may create confusion. Even where confusions do not occur, the English counting system does not give as strong clues to the base ten system, as do the counting systems of Chinese and other East Asian languages. This may be important to numerical development for several reasons. It may be easier to learn and remember the counting sequence if it is based on consistent and regular patterns (Miller et al., 1995). It may be easier to understand place value in written arithmetic if it corresponds closely to the oral counting system (Miura et al., 1988). More generally, an oral counting system that is both regular in itself, and transparently related to the written number system may contribute to the precision and accessibility of cognitive representations of number. This idea is a feature of several models of numerical cognition and how it may be influenced by the counting system (Nickerson, 1988; Zhang and Norman, 1995; Zhang and Wang, 2005; Bender and Beller, 2018). Most of these models focus mainly on adult numerical processing, but cross-cultural studies of children have provided some evidence for them.

Some evidence for superior understanding of base 10 and place value in children with highly transparent counting systems comes from work by Miura et al. (1988, 1993), Miura and Okamoto (2003), and Okamoto (2015). They initially investigated

the base 10 knowledge on Japanese (transparent counting system) and American (opaque counting system) first graders using Base 10 blocks. These blocks include unit blocks and tens blocks, with the tens blocks having ten segments of units shown on them. The studies revealed that Japanese children were significantly more likely to represent two-digit numbers using a combination of tens blocks and unit blocks, while the American children were more likely just to use unit blocks. This was interpreted as evidence that a transparent counting system facilitates understanding of the semantics of multi-digit numbers by using base-10 knowledge. Follow-up studies were done on different countries and confirmed this difference between the users of transparent and opaque counting systems and (e.g., Miura et al., 1988; Miura and Okamoto, 2003).

One problem with international comparisons is that children in different countries will differ with regard to a wide variety of educational and cultural influences: not just those involving language (Saxton and Towse, 1998). Studies of different language groups in the same country have suggested that language probably affects some specific numerical abilities, but not arithmetic globally. In Wales, most children study in English as elsewhere in the United Kingdom, but about 20% attend Welsh medium schools, where they use the transparent Welsh counting system for arithmetic. However, all schools in Wales follow the same national curriculum. Dowker et al. (2008) investigated the numerical abilities of 6- and year-old children attending English and Welsh medium schools in Wales. They found that there was no difference between the children at the English and Welsh medium schools regarding overall arithmetical performance, but that those in the Welsh medium schools were significantly better at reading and comparing two-digit numbers. Mark and Dowker (2015) studied children in Chinese and English medium schools in Hong Kong. They found that those in the Chinese medium school did perform somewhat better at a standardized arithmetic test, and at backward and forward counting, but, in contrast to the Welsh study, only younger children (6 to 7) and not older children (8 to 9) showed group differences in reading and comparing two-digit numbers.

The superior performance of speakers of languages with regular counting systems on some numerical tasks has led to the question of whether their internal spatial representations of numbers may be more precise. Most commonly, this is studied by means of number line estimation tasks. Number line estimation tasks ask participants to indicate an approximate position of a target number within a fixed range on a number line. Siegler and Booth (2004) found that performance on such tasks is related to performance on other numerical tasks, and that it improves with age. Not surprisingly, children find number lines that include a higher number range more difficult than those that involve a relatively low number range: Siegler and Booth (2004) found that they perform better on a 0–10 number line than a 0–20 number line, which is in turn easier for them than a 0–100 number line, while a 1 to 1000 number line is more difficult than any of the previous ones.

Some studies suggest that children using transparent counting systems are better at number line tasks than those using

more opaque counting systems, but results are conflicting. Siegler and Mu (2008) found that Chinese kindergarten children performed better than their American counterparts on mental number line estimation tasks involving a number line spanning from 1 to 100. Laski and Yu (2014) found that both Chinese and Chinese-American children performed better than monolingual English-speaking American children, but that children in China performed better on these tasks than Chinese-American children, suggesting that both linguistic *and* educational factors were important. By contrast, Muldoon et al. (2011) did not find such a difference between Chinese and Scottish 4- and 5-year-olds; and indeed when smaller number lines from 0 to 10 and 0 to 20 were included, the Scottish children performed better. This was despite the fact that the Chinese children did do better than the Scottish children on a standardized arithmetic test. Dowker and Roberts (2015) studied children in English and Welsh in Wales, and found a trend for children in Welsh medium schools to be more accurate and quicker on number line tasks, but the difference did not reach significance. However, the Welsh medium children *did* show significantly lower standard deviations than the English medium pupils, indicating more consistency and lower variability in performance.

There are also studies of children, who use counting systems that are *more* opaque than English, such as German, where the oral counting words are systematically inverted with respect to the written counting system, e.g., ‘drei und zwanzig’ (three and twenty) for 23. This might increase the potential for confusing tens and units when translating between the oral and written numbers systems. Such studies have indicated that children who use such counting systems are less accurate in placing numbers on empty number lines than children who use counting systems with little or no inversion (e.g., Helmreich et al., 2011; Krinzinger et al., 2011; Klein et al., 2013; Moeller et al., 2015; Bahnmueller et al., 2018a, in press).

The present study focuses on differences between English and Chinese-speaking children. There have already been a number of studies comparing numerical performance between these two language groups, as discussed earlier in the introduction. However, such studies have typically investigated either arithmetical performance *or* tasks involving numeral magnitudes *or* number line tasks. It is important to combine arithmetic tests with numeral magnitude and number line tasks, in order to investigate whether Chinese and English speaking children differ in a similar way for all of these tasks, or whether there are some tasks that favor Chinese-speaking children and some that do not. The key aim of the present study was to investigate and compare Chinese and English children’s numerical abilities on all these tasks.

A secondary aim was to look at specific aspects of the tasks that might influence whether, and to what extent, differences are found between Chinese and English children. For example, it is possible that there might be different results for tasks emphasizing different types of symbolic number representation. There are two main types of symbolic number representation that children use: number words and numerals. The numeral notations are transparent and regular in both languages. The number words are much more regular and predictable in

Chinese than in English, and as a consequence are also more transparently related to the numerals. One might therefore expect that English children would be mainly disadvantaged in tasks relating to number words: e.g., fast recognition of spoken number words, transcription of number words into numerical notation, and to some degree mental arithmetic. The disadvantage would be expected to be less pronounced in tasks based on numeral notations, such as written arithmetic and symbolic number comparisons. However, this would only be the case if there is a dissociation between representations of numerals and number words. As number word irregularities also decrease the relationship between number words and numerals, they could still affect numeral-based tasks if numeral-based tasks depend in part on translation from number words, or if the two interact.

In this study, we aimed to investigate Chinese- and English-speaking children’s arithmetical abilities. We gave the children a standardized arithmetic test, to check for global differences in arithmetical ability. We also gave them two tasks to measure more specific numerical abilities, which have sometimes been proposed to differ between users of transparent and opaque counting systems. One of these was a two-digit number comparison task, measuring the understanding of place value (Donlan and Gourlay, 1999; Dowker et al., 2008). The other was a task involving placement of visually- presented numbers on empty number lines of different range (Siegler and Booth, 2004; Moeller et al., 2009; Helmreich et al., 2011; Link et al., 2014; Schneider et al., 2018, in press). Both symbolic number comparison (Göbel et al., 2014) and number line task performance (Petitto, 1990; Schneider et al., 2009; Schneider et al., 2018, in press) been found to predict current and future arithmetical performance. Sasanguie et al. (2013) found that both symbolic number comparison and number line task performance in 6-to-8-year-olds predicted their future arithmetical performance, though symbolic number comparison was the strongest predictor. Schneider et al. (2018, in press) carried out a meta-analysis, which also indicated that both symbolic number comparison and number line task performance predicted arithmetical performance, but suggested that number line task performance was the strongest predictor in 6-to-9-year-olds, and that the two types of task were equally strong predictors in older children.

We predicted that the Chinese pupils would perform better in the standardized arithmetic test, on the basis that in general, Chinese pupils perform better than English pupils in most comparisons of mathematical performance, and in particular, Mark and Dowker (2015) found that Chinese pupils performed better than English pupils on the same standardized arithmetic test.

We predicted more tentatively that they would do better on the number comparison task, as this had been found for Welsh- versus English-speaking children (Dowker et al., 2008), and Chinese versus German children (Lonneman et al., 2016), though not in Mark and Dowker’s (2015) study of Chinese-speaking versus English-speaking children. There is also evidence that performance on two-digit number comparison tasks is affected by other aspects of counting systems, such as the inversion property

of some languages including German (Nuerk et al., 2005) and the vigesimal structure of numbers over 60 in French (Van Rinsveld and Schiltz, 2016).

In addition, we predicted that English children and Chinese children might be differentially affected by the difficulty of the comparison type. Following Donlan and Gourlay (1999), the number comparison task included three different types of comparison. Transparent comparisons were those that did not involve any conflict between the relative values of the decades and the units. Either the numbers shared a unit value lower than either decade value (e.g., 21 vs. 71), or both comparisons contained repeated digits (e.g., 33 vs. 88). Misleading comparisons were those where the unit values differed in the opposite direction from the decade values: e.g., 32 vs. 29. Reversible comparisons were similar to misleading comparisons, but specifically involved pairs where each number reversed the decade and unit values of the other: e.g., 91 vs. 19. We predicted that the Chinese and English children would show greater differences in speed and accuracy for comparisons involving misleading or reversible comparisons than for transparent comparisons. This was because, if Chinese children have more solid representations of two-digit numbers, they would be better able to focus just on comparing the tens and to resist interference from the relative values of the units; and that this would show up mainly in comparisons where there is a conflict between the relative values of the decade numbers and the unit numbers.

We also predicted that the Chinese-speaking children would be more accurate and faster on the number line tasks, due to a greater understanding of, and facility with, multi-digit numbers and their relative magnitudes. While there have been a number of studies of Chinese children's number line task performance, for the most part such studies have not, to our knowledge, examined differences between different number line ranges, with the notable exception of Siegler and Mu's (2008) study, and that only looked at preschoolers with limited experience of the larger number line ranges. This made it important to investigate whether number line range had similar or different effects on Chinese and English children. We predicted that *both* groups of children would find the number lines with larger number ranges would be more difficult for than those with smaller number ranges, and would thus show lower accuracy scores and higher reaction times for the number lines with the larger ranges. However, we also predicted that the differences between Chinese and English-speaking children would be greater for number lines with ranges of 0–20 or more than for the 0–10 number line, because the greater transparency of the Chinese counting system only comes into play for numbers over 10. Thus, any advantages to children of using the more transparent Chinese counting system would be expected to emerge only at the point where their counting system *does* become more transparent than the English counting system.

Thus, we expected that combining the standardized arithmetic test, the number comparison task and the number line task would shed light on what aspects of numerical processing are most influenced in this age group by cultural differences, and on whether any such differences are readily explainable in terms of

differences in internal representation of numbers, or are better explained in other ways.

MATERIALS AND METHODS

Participants

Seventy five children (30 girls) participated in the study. They were tested at the end of the first term of their second year of primary school. They included 35 children (10 girls, 25 boys) attending primary schools in Oxford, and 40 (20 girls, 20 boys) attending primary schools in Hong Kong. The mean age of the children was 7.2 years (*SD*: 0.77). The English children had a mean age of 7.09 years (*SD*: 0.95) and the Chinese children 7.3 years (*SD*: 0.56). There was no significant age difference between the two language groups, as confirmed by an independent-samples *t*-test [$t(73) = 1.204$; $p = 0.56$; Cohen's $d = 0.26$]. All of the Oxford children spoke English as their first language, and none had any knowledge of Chinese or any other East Asian language. All of the Hong Kong children spoke Cantonese as their first language. Most had had some limited exposure to English, but all were taught their main school subjects, including mathematics, in Chinese.

Ethical approval for the study was obtained from Oxford University's Central University Research Ethics Committee; and informed written parental consent was obtained for all participants.

Procedure

All participants were given the same tests in the same order: the standardized arithmetic test, followed by the number comparison test, followed by the number line estimation test. The standardized arithmetic test was completed with pencil and paper, and the other tests were given on a Lenovo G50 laptop. Instructions were given to the children in their native language by a bilingual experimenter. Participants were tested individually in a quiet room in their schools. The whole testing session lasted for approximately 40 min.

Standardized Arithmetic Test

Participants were given the British Abilities Scales (BAS) 2nd edition Basic Number Skills test (Elliott et al., 1996), designed to assess the numerical abilities of children between the ages of 6 and 16. The assessment consists of a series of questions, split into different sections which increase in difficulty as the children progress. Most of the questions involve written calculation. All of the four arithmetical operations are included. There are 46 items in total, arranged in six blocks (A to F); the first four blocks consist of eight items each, and the last two blocks have seven items each. The test is stopped when the child makes four or more errors within a section. In practice, no child progressed further than Section D.

The first section, Section A, includes four numbers that children are asked to read aloud: 100, 12, 40, and 31. It also includes four written arithmetic problems, presented in vertical form: $2 + 3$, $4 - 1$, $9 + 5$, and $18 - 5$. The second section, Section B, includes two numbers that children are asked to read

aloud: 215 and 370. It also includes a request to point out the orally presented number ‘five hundred and ninety four’ out of five written numbers ranging from 54 to 50094. It includes five written arithmetic problems, presented in vertical form: $15 + 23$; 2×4 ; $17 - 5$; $13 + 99$; and $38 + 57$.

The third section, Section C, includes eight written arithmetic problems: two involving multiplication of a two-digit number by a single-digit number; three involving division of two-digit numbers by single-digit numbers; two two-digit subtractions requiring borrowing; and one involving addition of decimals ($45.01 + 57.89$).

The fourth section, Section D, includes eight written arithmetic problems. These include one problem involving addition of fractions ($1/8 + 1/4$); one subtraction of fractions ($2/3 - 1/3$); two problems involving writing decimals as percentages; one problem involving division of a two-digit number by a smaller two-digit number; one problem involving division of a three-digit number by a two-digit number; one involving multiplication of two two-digit numbers; and one involving decimal arithmetic ($3(2.7 + 9.3)$).

The fifth and sixth sections, Sections E and F, will not be described as no child reached these sect.

Number Comparison Test

Children were given Donlan and Gourlay (1999) number comparison test. The task was slightly modified in order for it to be used with current computer systems.

There were three types of number pair stimuli – transparent, misleading and reversible. Transparent stimuli were defined as number pairs that varied in the number of tens but had the identical number of units (e.g., 91 and 71) or with repeated digits (e.g., 55 and 44). Participants could make judgments for the response by only looking at the tens. Misleading stimuli are number pairs with a higher number of digits in the smaller two-digit number than that in the bigger one (e.g., 31 and 27). Judgments to these stimuli require holistic processing of both tens and units for correct comparison. Reversible pairs included number pairs with opposite tens and digit positions (e.g., 64 and 46). The items were presented in a new random order for each participant, and were not presented in blocks. Error scores and reaction times were the main measures of the task. The full set of stimuli is displayed in **Table 1**.

TABLE 1 | Number comparison task (adapted from Dowker et al., 2008).

Transparent	Misleading	Reversible
73 43	51 47	85 58
66 55	61 18	61 16
54 94	53 39	14 41
70 10	62 59	56 65
96 86	27 42	43 34
11 99	71 91	76 67
60 50	17 51	39 93
71 91	43 38	25 52
44 55	27 31	46 64

EPrime 2.0 was used to present pairs of numbers side by side on the laptop. The viewing distance was 60 cm. The presentation sequence consisted of a fixation point (500 ms), followed by a slide with two two-digit numbers presented side by side, of approximately 5 cm apart. The number slides only changed to a fixation screen after the laptop detected a response. The whole process was repeated for the remaining trials.

When pairs of two-digit numbers were presented on each slide, participants were asked to give responses on the keyboard by indicating whether the number on the left or the number on the right was the larger number. Before starting, participants were instructed to give responses by two keys on either ends of the keyboard, and their response (left or right key) was consistent to their opinions of where the larger numbers were (left or right key respectively). To prevent contradiction with comparing the physical sizes of stimuli, subjects were given three practice trials to familiarize themselves with the equally-sized numbers.

Number Line Estimation Test

The children were given four number line estimation tasks (0–10, 0–20, 0–100, and 0–1000) in that order. The test was based on that used by Siegler and Booth (2004), and the sets were those used by Moeller et al. (2011). Once again, in this study, the tasks were carried out on a laptop screen, with the program devised using EPrime 2.0. The number line was presented, at the bottom of the screen, as a long green horizontal rectangle of length 16.8 cm and width 2.4 cm. The ends of the number line were clearly shown (font size: 70) on both sides of the rectangle – 0 on the left; 10/20/100/1000 on the right, depending on the task. Target numbers were presented visually at the center of the screen (font size: 80) one at a time. Before the start of the test, each child was given a couple of practice trials in which they were asked to point to the approximate positions of 5 and 8 on a 0–10 number line. The aim was to check if the children understood the meanings of 0 and 10 at either ends of the rectangle. If the participant demonstrated that they understood that 8 was on the right of 5, the experimenter said, ‘Well done. Now let’s get on to the real thing.’ All children used the pointer of a mouse to give response by clicking on the various positions on the rectangle. The rectangle was designed to appear continuous, but was segregated into 100 slices after a response was given. The respective rectangle that was hit was recorded as a percentage response on the number line. The main measure of the tasks was the percentage difference between the response value on the rectangle and target number (percentage absolute error; henceforth abbreviated to PAE). After each response, there was a 1000ms-delay. Responses made outside the area of the rectangle were not detected by the program, and therefore the child would be reminded to respond again inside the rectangle. There were 10 trials each for 0–10 and 0–20 tasks, and 19 trials each for 0–100 and 0–1000 tasks. On each number line, the order of the target numbers to be estimated was randomized across all children.

The 10 target numbers on the 0–10 number line were 6, 0, 7, 2, 8, 1, 4, 9, and 3. The 10 target numbers on the 0–20 number line were 10, 12, 1, 13, 4, 15, 19, 7, 27, and 5. The 19 target numbers

on the 0–100 number line were 50, 27, 2, 64, 35, 7, 13, 99, 75, 47, 3, 11, 82, 95, 9, 17, 6, 18, and 53. The 19 target numbers were 500, 4, 96, 465, 287, 989, 26, 432, 173, 823, 87, 124, 367, 679, 57, 107, 73, 92, and 725.

There was no set time limit, but children were asked to respond as quickly as possible, overt use of strategies other than estimation (such as counting) was discouraged. The scoring measures used were Percentage of Absolute Error (PAE), and also reaction time, as used e.g., by Schneider et al. (2009).

RESULTS

The data were entered and analyzed using SPSS Statistics 22 (SPSS, Inc. 2013).

Group Comparisons

Standardized Arithmetic Test

The mean raw score on the arithmetic test was 16.4 ($SD = 4.6$). The Chinese children obtained a mean score of 18.35 ($SD: 3.51$). The English children obtained a mean score of 14.17 ($SD = 4.72$). An independent samples t -test showed that this difference was significant [$t(73) = 4.39$; $p < 0.01$; Cohen's $d = 1.02$], with the Chinese children performing significantly better than the English children.

Number Comparison Test

A repeated-measures ANOVA was carried out with Comparison type (Easy vs. Misleading vs. Reversible) as the within-participants variable Language Group (English versus Chinese) as the grouping factors; and Number Comparison score as the dependent variable. Though there was a trend toward greater accuracy by Chinese than English children, the group difference did not reach significance [$F(1,73) = 2.86$; $p = 0.068$; $\eta_p^2 = 0.209$]. There was no significant within-participants effect [$F(2,146) = 1.075$; $p = 0.303$; $\eta_p^2 = 0.01$], nor any significant interaction effect between Language Group and Number Comparison score [$F(2,146) = 0.8$; $p = 0.449$; $\eta_p^2 = 0.011$].

Another repeated-measures ANOVA was carried out with Comparison type (Easy vs. Misleading vs. Reversible) as the within-participants variable; Language Group (English versus Chinese) as the grouping factors; and Reaction Time score in milliseconds as the dependent variable. There was a strong between-participants effect of Language Group [$F(1,73) = 50.374$; $p < 0.001$; $\eta_p^2 = 0.415$]. Chinese children were much faster than English children. There was also a significant within-participants effect [$F(2,146) = 7.352$; $p = 0.001$; $\eta_p^2 = 0.094$]. Pairwise comparisons showed that reaction times were longer for Misleading than Reversible problems, and for both Misleading and Reversible problems than Easy problems. There was, however, no significant interaction effect between Language Group and Comparison type [$F(2,146) = 0.95$; $p = 0.389$; $\eta_p^2 = 0.013$]. Thus, the language groups differed in overall performance, but not with regard to the relative difficulty of the comparison types.

Number Line Tasks

The Number Line Mean Reaction Times in milliseconds are also given in **Table 2**. There was a significant within-participants effect of Number Line Range [$F(3,219) = 15.114$, $p < 0.001$; $\eta_p^2 = 0.186$]. Pairwise comparisons showed no significant difference in Mean Reaction Time between the 0–20 and the 0–1000 number lines ($p = 0.47$) and only a trend toward significance between the 0–100 and 0–1000 number lines ($p = 0.084$), but all other differences between number lines were significant. The difference between the 0–20 and the 0–100 number lines reached significance ($p = 0.031$) and the differences between the 0–10 and the 0–20 number lines; the 0–10 and the 0–100 number lines; and the 0–10 and the 0–1000 number lines were all highly significant ($p < 0.001$). There was a significant between-participants effect of Language Group [$F(1,73) = 12.69$; $p < 0.001$; $\eta_p^2 = 0.161$]. However, there was no significant interaction between Language Group and Number Line Range [$F(3,219) = 1.28$; $p = 0.283$; $\eta_p^2 = 0.161$].

For each participant the mean PAE score for each number line was calculated. The PAE score of each trial was the absolute distance between the true position of the target number and the response. **Table 3** gives the mean PAE and reaction time for each number line in each language group. Two repeated-measures ANOVAs were conducted with Number Line Range (0–10 vs. 0–20 vs. 0–100 vs. 0–1000) as the within-participants factor; Language Group (English versus Chinese) as the grouping factors, and Mean PAE and Mean Reaction Time as dependent variables.

For Mean PAE, there was a significant within-participants effect of Number Line Range [$F(3,219) = 68.06$; $\eta_p^2 = 0.49$]. Pairwise comparisons showed no significant difference in Mean PAE between the 0–10 and 0–100 number lines ($p = 0.15$) but all other comparisons were significant. The mean difference in PAE between the 0–10 and the 0–20 number line reached significance ($p = 0.031$) and the differences between the 0–10 and the 0–1000 number lines; the 0–20 and the 0–100 number lines, the 0–20 and the 0–1000 number lines; and the 0–100 and the 0–1000

TABLE 2 | Mean scores (out of 9) and RTs (in milliseconds) by each language group on easy (transparent), misleading, and reversible number comparison items.

Number comparison item type	Language group		
	Chinese ($N = 40$)	English ($N = 35$)	Combined ($N = 75$)
Scores			
Easy	8.75 (0.49)	8.71 (0.67)	8.73 (0.58)
Misleading	8.68 (0.57)	8.69 (0.58)	8.68 (0.57)
Reversible	8.95 (0.22)	8.77 (0.49)	8.87 (0.39)
Total score*	26.38 (0.90)	26.12 (1.00)	26.26 (1.04)
RTs			
Easy	1512.25 (68.43)	3443.9 (1816.32)	2285.46 (1613.64)
Misleading	1657.78 (678.4)	3775.83 (1838.64)	2606.21 (1690.46)
Reversible	1697.71 (650.28)	3885.9 (1704.23)	2664.29 (1633.52)
All problems	1622.58 (612.47)	3678.54 (1705.68)	2551.99 (1599.31)

RT, reaction time. Standard deviations are given in brackets. *Total score is out of 27.

TABLE 3 | Mean percentages of absolute error (PAE) and RTs (in milliseconds) by each language group on the number line tasks.

Number line	Language group		
	Chinese (<i>N</i> = 40)	English (<i>N</i> = 35)	Combined (<i>N</i> = 75)
PAE			
0–10	7.1 (3.67)	8.21 (4.13)	7.61 (3.90)
0–20	7.23 (4.34)	5.68 (2.48)	6.52 (3.67)
0–100	9.25 (5.22)	9.78 (6.61)	9.5 (5.86)
0–1000	17.3 (9.12)	16.6 (11.15)	16.98 (10.04)
RTs			
0–10	5060.05 (1893.28)	6803.44 (3091.47)	5829.19 (2622.35)
0–20	4204.6 (1746.47)	5233.31 (1701.35)	4658.44 (1728.58)
0–100	3795.68 (1745.45)	4912.48 (2035.38)	4289.39 (1946.19)
0–1000	3890.19 (1348.24)	5663.88 (2734.88)	4672.7 (2242.44)

RT, reaction time. Standard deviations are given in brackets.

number lines were all highly significant ($p < 0.001$). There was no significant effect of Language Group [$F(1,73) = 0.021$; $p = 0.895$; $\eta_p^2 = 0$]. Nor was there any significant interaction between Language Group and Number Line Range [$F(3,219) = 0.899$; $p = 0.443$; $\eta_p^2 = 0.012$].

DISCUSSION

Overall, the results supported the hypotheses that Chinese children would perform better on tests of numerical abilities, but this varied to some degree with the measures used. The Chinese children performed better on a standardized arithmetic test. They were faster but not more accurate on a number comparison task; though near-ceiling effects might have contributed to the lack of group differences in accuracy. They had significantly faster reaction times to the number line tasks, but did not differ significantly in accuracy, which in this task cannot be attributed to ceiling effects.

The better performance by Chinese than English children in the standardized arithmetic test was in line with our predictions, and similar to findings in many other studies (e.g., Mark and Dowker). This is likely to be due to several factors, which may include the transparency of the counting system; the greater length of time devoted to arithmetic in Chinese schools; stronger societal value placed on mathematics in China; and possibly differences in teaching methods. The superior performance by Chinese children is particularly striking in view of the fact that the test was developed and standardized in Britain, making it very unlikely that Chinese children would have had any direct preparation for it.

The prediction that the Chinese children would do better than the English children in number comparison tasks was partially supported. They were faster, but did not differ in accuracy. Their faster reaction times give some support to Miura et al. (1988) hypothesis that the use of transparent counting systems may improve understanding of the semantics of the base ten system, and are in line with Dowker et al. (2008) findings comparing English and Welsh children, and Lonneman et al. (2016) findings

comparing Chinese and German children. This result suggests that certain multi-digit number tasks are indeed easier for children who speak languages with transparent counting systems. The lack of group differences in *accuracy* may not go against this hypothesis, given the near-ceiling effects for accuracy, mentioned above; and also because of the possibility of a speed-accuracy trade-off. However, the results do not confirm the prediction that there would be an interaction between group and comparison type, and thus do not support a view that the Chinese and English children are likely to be using fundamentally different strategies, or to have fundamentally different number representations. Both groups were faster at comparing easy (transparent) pairs than misleading pairs, with reversible pairs coming in between. The fact that the reversible pairs were somewhat easier than the other misleading pairs may be due to the fact that fewer numbers needed to be kept in working memory. However, the difference was not great: the misleading and reversible pairs were more similar to one another than they were to the transparent pairs, supporting earlier findings by Nuerk et al. (2005). Contrary to the predictions, English children were not more affected than the Chinese children by the comparison type.

The results also give partial, but not total, support for the hypothesis that children, who use a transparent counting system, would be better at placing numbers on an internal number line. Once again, the Chinese children were faster, but they were not more accurate. Again, a speed-accuracy trade-off may have contributed to the results. It should be noted that in this case, different cultural influences may have been in conflict. The Chinese children had a more transparent counting system, and may also have been subject to other positive educational influences; but the English children had more specific experience with number lines.

Number lines play a significant part in United Kingdom mathematics instruction. The United Kingdom national curriculum for primary school mathematics¹ indicates that pupils are expected to be taught to use number lines throughout years 1 to 6, with increasing levels of sophistication. This in part related to an emphasis in the United Kingdom on mathematical estimation in general. On the other hand, a careful scan of the HK's primary school curriculum reveals no mention of either 'number estimation' or 'number line'². The focus of teaching in HK appears to be more geared toward instruction in procedures for exact mental and written calculation. Although systematic quantitative data are still needed, brief interviews with the children indicated that the United Kingdom pupils had had practice with the use of number lines at school, while most Hong Kong pupils reported a lack of experience with them. The Hong Kong pupils tended to respond to the number line tasks by utilizing strategies for counting exact quantities by trying to visualize imaginary counters on the stimulus, without taking

¹National curriculum in England: framework for key stages 1 to 4 (effective from 1 September, 2015 to 31 August, 2016) – <https://www.gov.uk/government/publications/national-curriculum-in-england-framework-for-key-stages-1-to-4/the-national-curriculum-in-england-framework-for-key-stages-1-to-4>.

²Contents of Curriculum, Learning Targets of Key Stages 1 and 2 – http://www.edb.gov.hk/attachment/en/curriculum-development/kla/ma/curr/chapter%204_1.pdf.

much notice of the extremes of the number line; and verbalized counting far more than the United Kingdom pupils did. This was inferred from consistent patterns of verbalization of counting in the HK sample but not in the United Kingdom sample.

The number line range had significant effects on performance by both groups, supporting findings by Siegler and Booth (2004) and others. Number lines with larger ranges were generally more difficult, in that they elicited larger errors. There was little difference between performance on the 1–10 and the 1–20 number line, but the PAE increased with increasing number line range beyond 20. Reaction time on the other hand actually decreased from the 1 to 10 number line to those with higher ranges, though this effect showed signs of reversal for the number line with a 1–1000 range. This may be in part due to practice effects, as the 1 to 10 number line was given first, and possibly fatigue on the 1–1000 number line. It may also reflect changes in strategy, with a reduction in counting-related strategies as the number line range increased.

The fact that there was no interaction between group and number line range, with regard to either accuracy or reaction time, suggests that the English and Chinese children were not using fundamentally different strategies for the number line tasks. It would be desirable in future studies to investigate and compare the strategies of English and Chinese children directly, perhaps combining the strategy analyses of Link et al. (2014) with the eye tracking measures developed by Schneider et al. (2018).

Thus, the study supports the view that the transparency of a counting system influences some but not all numerical abilities. It is important to remember that the counting system is by no means the only reason for cultural and national differences in mathematics. As already mentioned, such differences are influenced by educational methods and by cultural attitudes to education. When children, who use different counting systems, receive the same curriculum, they tend to perform similarly on arithmetic tests, though often differing in more specific numerical abilities (Dowker et al., 2008; Dowker and Roberts, 2015). Thus, it is most likely that the differences in performance on the arithmetic test in the present study were due to educational and/or other cultural factors, while the differences on other numerical abilities may more likely to have been due to linguistic factors.

There is a caveat to be made here: since the group differences were found for reaction time but not for accuracy, it is possible that they reflect differences in speed of responses to tasks in general, rather than numerical tasks in particular. Chinese children may either have a generally faster speed of processing, or be more likely to interpret test situations as requiring speedy responses. Because of a possible speed-accuracy trade-off, a greater Chinese emphasis on speed could have led to an underestimation of differences in ability to produce accurate responses. Future studies should include non-numerical control tasks, to check for this possibility. Also, even if the effects are based on the counting system, they might reflect not the greater transparency of the Chinese counting system, but the relative shortness and faster pronounceability of Chinese number words (Ellis and Hennelly, 1980). This possibility could be tested in the future by making direct comparisons between Chinese- and

Welsh-speaking children, as their counting systems are similarly transparent, but Welsh number words are *longer* than English number words.

Further studies are needed to investigate the relative importance of linguistic, educational and other cultural factors. Such studies should if possible include investigations of specific educational content, such as the use of number lines, and cultural factors such as differences in finger counting techniques (Göbel et al., 2011). Also, future studies should incorporate larger samples with a wider variety of ages, languages and backgrounds. One potential limitation of the present study is that there was relatively limited information about possible social and economic differences between the groups. The backgrounds appeared to be similar (varied but predominantly middle-class); but quantitative information on this matter was not collected. This should be investigated more systematically in future research.

Most research on cross-linguistic effects on arithmetic has focused on the effects of counting system transparency. The present study has combined investigation of standardized test performance with investigation of more basic numerical abilities, and indicates that counting system transparency does indeed have some effect on both. Future studies should now look more at other linguistic differences that might affect arithmetic and number processing (Göbel et al., 2011; Dowker and Nuerk, 2016; Bahnueller et al., 2018b). These include, for example, phonological factors such as the length and pronunciation speed for number words; grammatical factors such as whether a language has dual and plural markers; and semantic factors such as the ways in which numerical concepts such as ‘few,’ ‘many,’ ‘more,’ and ‘less’ are represented in words and symbols.

Future studies should also include measures of domain-general factors, such as IQ, working memory, and verbal and spatial ability, which could directly influence arithmetical and numerical abilities, and possibly also mediate or moderate relationships between numerical abilities and arithmetic. Research is providing increasing evidence for the role played by such domain-general factors in numerical development (Schneider et al., 2008; Schneider et al., 2018, in press). For example, Simms et al. (2016) have found that visuospatial and visuomotor abilities explain much of the relationships between number line task performance and arithmetic in 8- to 10-year-olds; though they also found that PAE (unlike some other number line performance measures) predicted arithmetic even after controlling for visuomotor and visuospatial abilities. Other researchers have found that number line performance is correlated with domain-general spatial abilities (Gunderson et al., 2012); measurement skills (Cohen and Sarnecka, 2014) and overall IQ (Schneider et al., 2009). It is important to investigate whether these and other domain-general abilities show similar relationships to numerical abilities in different language groups.

A potential limitation is that the tasks, including the number line tasks, were given in a fixed order. This was done, so as to avoid the need to use presentation order as an additional variable; but it makes it difficult to draw conclusions as to whether the lower reaction times to lines with higher number ranges were due to practice effects or to strategy changes. Future studies should look at whether there are order effects.

The present study adds to our knowledge base about cultural differences in numerical abilities, by demonstrating that Chinese and English children do indeed appear to show differences in numerical tasks as well as in formal arithmetic. The Chinese children were much more accurate than the English children in the formal arithmetic test. They did not show such differences in accuracy in the non-arithmetical numerical abilities. However, they did show striking differences in *speed*: the Chinese children were much faster than the English children on both the number comparison task and the number line task.

The results do not support the study's secondary prediction that the differences would affect tasks involving number words but not those involving numerical notation. The number line tasks involved numbers presented *only* in numerical notation, and not through spoken words; and yet group differences were found. This suggests that, at least with children at this age, numerical notations and number words may not be processed totally independently. However, we need to be cautious in making strong interpretations of these results, since the main purpose of the study was not to compare numerical notations with spoken words, and they were not systematically varied.

One possible reason for the findings that group differences were stronger for arithmetic than for accuracy (though not speed) on non-arithmetical tasks is that the arithmetic problems might have relied more on verbal processes, while the number line and number comparison problems might have relied more on visuospatial processes. The transparency of the verbal counting system would be likely to have a greater effect on verbal than visuospatial processes. To solve arithmetic problems, the children might have relied at least partially on verbal processes that might account for the differences between groups. Verbal processes might have been slightly more efficient with more transparent verbal number words (i.e., Chinese). On the contrary, number lines would be rather tap into visuo-spatial processes and an internal number representation without any need of verbal processing and, by consequence, produce reduced differences between the groups. In brief, the differences between the groups might emerge when the numerical tasks involve number words at the processing level (even though the task material itself is not presented in a verbal format), such as arithmetic typically.

There are numerous educational and cultural differences between Chinese and English children that are likely to contribute to the results. It is, however, likely that the counting system is a significant contributory factor, as some other studies have found differences between users of transparent and non-transparent counting systems even within the same geographical region and educational system (Dowker et al., 2008; Mark and Dowker, 2015) and even between performance by the same individuals using different counting systems within the Czech language (Pixner et al., 2011). The results, however, do not indicate that Chinese and English children have fundamentally different internal representations of number, though this may depend on age, and findings might be different for older or younger children. It is perhaps more likely that a transparent counting system facilitates arithmetical and numerical performance by making the numerical characteristics of, and the relationships and differences between, two-digit numbers more salient, and by reducing the load that multi-digit numbers place on working memory.

ETHICS STATEMENT

The study was carried out in accordance with the guidelines of the Central University Research Ethics Committee of Oxford University. As the study involved work with children, written parental consent was obtained for all participants, using an opt-in procedure, where parents were given an information sheet and signed a consent form. All aspects of the study were carried out in accordance with the university's Inter-Divisional Research Ethics Committee's Protocol 25, which sets out the expected procedures for work with children in schools. The project was approved by the Central University Research Ethics Committee of Oxford University.

AUTHOR CONTRIBUTIONS

AD provided the tests and took the main role in writing this article. AML carried out the experiments and did all necessary translations. AD and AML worked together in designing the project and carrying out the statistical analyses.

REFERENCES

- Askew, M., Hodgen, J., Hossain, S., and Bretscher, N. (2010). *Values and Variables: Mathematics Education in High-Performing Countries*. London: Nuffield Foundation.
- Bahnmueller, J., Maier, C., Goebel, S., and Moeller, K. (2018a). Direct evidence for linguistic influences in two-digit number processing. *J. Exp. Psychol. Learn. Mem. Cogn.* doi: 10.1037/xlm0000642 [Epub ahead of print].
- Bahnmueller, J., Nuerk, H. C., and Moeller, K. (2018b). A taxonomy proposal for types of interactions of language and place-value processing in multi-digit numbers. *Front. Psychol.* 9:1024. doi: 10.3389/fpsyg.2018.01024
- Bender, A., and Beller, S. (2018). "Numeration systems as cultural tools for numerical cognition," in *Mathematical Cognition and Learning: Language and Culture in Mathematical Cognition*, eds D. B. Berch, D. C. Geary, and K. Mann Koepke (Cambridge, MA: Academic Press), 297–320. doi: 10.1016/B978-0-12-812574-8.00013-4
- Cohen, D. J., and Sarnecka, B. W. (2014). Children's number-line estimation shows development of measurement skills (not number representations). *Dev. Psychol.* 50, 1640–1652. doi: 10.1037/a0035901
- Donlan, C., and Gourlay, S. (1999). The importance of nonverbal skills in the acquisition of place-value knowledge: evidence from normally-developing and language-impaired children. *Br. J. Dev. Psychol.* 17, 1–19. doi: 10.1348/026151099165113
- Dowker, A., Bala, S., and Lloyd, D. (2008). Linguistic influences on mathematical development: how important is the transparency of the counting system. *Philos. Psychol.* 21, 523–538. doi: 10.1080/09515080802285511
- Dowker, A., and Nuerk, H.-C. (2016). Linguistic influences on mathematics. *Front. Psychol.* 7:1035. doi: 10.3389/fpsyg.2016.01035
- Dowker, A., and Roberts, M. (2015). Does the transparency of the counting system affect children's numerical abilities? *Front. Psychol.* 6:945. doi: 10.3389/fpsyg.2015.00945

- Elliott, G., Murray, D., and Pearson, L. (1996). *British Abilities Scales*, 2nd Edn. Windsor: National Foundation for Educational Research.
- Ellis, N. C., and Hennessey, R. A. (1980). A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculation in Welsh and English. *Br. J. Psychol.* 71, 43–52. doi: 10.1111/j.2044-8295.1980.tb02728.x
- Geary, D. C., Salthouse, T. A., Chen, G., and Fan, L. (1996). Are East Asian versus American differences in arithmetical ability a recent phenomenon? *Dev. Psychol.* 32, 254–262. doi: 10.1037/0012-1649.32.2.254
- Gibb, N. (2014). Rote learning will bring pupils 'on par with peers in Far East'. Available at: <https://www.telegraph.co.uk/education/educationopinion/10645954/Rote-learning-will-bring-pupils-on-par-with-peers-in-Far-East.html>
- Göbel, S. M., Shaki, S., and Fischer, M. H. (2011). The cultural number line: a review of cultural and linguistic influences on the development of number processing. *J. Cross Cult. Psychol.* 42, 543–565. doi: 10.1177/0022022111406251
- Göbel, S. M., Watson, S. E., Lervag, A., and Hulme, C. (2014). Children's arithmetic development: it is number knowledge, not the approximate number sense, that counts. *Psychol. Sci.* 25, 789–798. doi: 10.1177/0956797613516471
- Gunderson, E. A., Ramirez, G., Beilock, S. L., and Levine, S. C. (2012). The relation between spatial skill and early number knowledge: the role of the linear number line. *Dev. Psychol.* 48, 1229–1241. doi: 10.1037/a0027433
- Helmreich, I., Zuber, J., Pixner, S., Kaufmann, L., Nuerk, H.-C., and Moeller, K. (2011). Language effects on children's non-verbal number line estimations. *J. Cross Cult. Psychol.* 42, 598–613. doi: 10.1177/0022022111406026
- Hess, R. D., Chang, C.-M., and McDevitt, T. M. (1987). Cultural variations in family beliefs about children's performance in mathematics: comparisons among people's republic of China, Chinese-American, and Caucasian-American families. *J. Educ. Psychol.* 79, 179–188. doi: 10.1037/0022-0663.79.2.179
- Jerrim, J., and Vignoles, A. (2015). *The Causal Effect of East Asian 'Mastery' Teaching Methods on English Children's Mathematics Skills*. DoQSS Working Paper No. 15-05. London: University College London Institute of Education.
- Klein, E., Bahnmüller, J., Mann, A., Pixner, S., Kaufmann, L., Nuerk, H. C., et al. (2013). Language influences on numerical development—inversion effects on multi-digit number processing. *Front. Psychol.* 4:480. doi: 10.3389/fpsyg.2013.00480
- Krinzinger, H., Gregoire, J., Desoete, A., Kaufmann, L., Nuerk, H., and Willmes, K. (2011). Differential language effects on numerical skills in second grade. *J. Cross Cult. Psychol.* 42, 614–629. doi: 10.1177/0022022111406252
- Laski, E. V., and Yu, Q. Y. (2014). Number line estimation and mental addition: examining the potential roles of language and education. *J. Exp. Child Psychol.* 117, 29–44. doi: 10.1016/j.jecp.2013.08.007
- Link, T., Huber, S., Nuerk, H. C., and Moeller, K. (2014). Unbounding the mental number line—new evidence on children's spatial representation of numbers. *Front. Psychol.* 4:1021. doi: 10.3389/fpsyg.2013.01021
- Lonneman, J., Linkersdorfer, J., Hasselhorn, M., and Lindberg, S. (2016). Differences in arithmetical performance between Chinese and German children are accompanied by difference in processing of symbolic numerical magnitude. *Front. Psychol.* 7:1337. doi: 10.3389/fpsyg.2016.01337
- Ma, L. (1999). *Knowing and Teaching Elementary Mathematics: Teachers' Understanding of Mathematics in China and the United States*. Mahwah, NJ: Erlbaum.
- Mark, W., and Dowker, A. (2015). Linguistic influence on mathematical development is specific rather than pervasive: revisiting the Chinese number advantage in Chinese and English children. *Front. Psychol.* 6:203. doi: 10.3389/fpsyg.2015.00203
- Miller, K. F., Smith, C. M., Zhu, J., and Zhang, H. (1995). Preschool origins of cross-national differences in mathematical competence: the role of number-naming systems. *Psychol. Sci.* 6, 56–60. doi: 10.1111/j.1467-9280.1995.tb00305.x
- Miura, I. T., Kim, C. C., Chang, C. M., and Okamoto, Y. (1988). Effects of language characteristics on children's cognitive representation of numbers: cross-national comparisons. *Child Dev.* 59, 1445–1450. doi: 10.2307/1130659
- Miura, I. T., and Okamoto, Y. (2003). "Language supports for mathematics understanding and performance," in *The Development of Arithmetic Concepts and Skills: Constructing Adaptive Expertise*, eds A. J. Baroody and A. Dowker (Mahwah, NJ: Erlbaum), 229–242.
- Miura, I. T., Okamoto, Y., Kim, C. C., Steere, M., and Fayol, M. (1993). First graders' cognitive representation of number and understanding of place value: cross-national comparisons – France, Japan, Korea, Sweden, and the United States. *J. Educ. Psychol.* 85, 24–30. doi: 10.1037/0022-0663.85.1.24
- Moeller, K., Pixner, S., Kaufmann, L., and Nuerk, H. C. (2009). Children's early mental number line: Logarithmic or decomposed linear? *J. Exp. Child Psychol.* 103, 503–515. doi: 10.1016/j.jecp.2009.02.006
- Moeller, K., Pixner, S., Zuber, J., Kaufmann, L., and Nuerk, H.-C. (2011). Early place-value understanding as a precursor for later arithmetic performance – a longitudinal study on numerical development. *Res. Dev. Disabil.* 32, 1837–1851. doi: 10.1016/j.ridd.2011.03.012
- Moeller, K., Shaki, S., Goebel, S. M., and Nuerk, H.-C. (2015). Language influences number processing – a quadrilingual study. *Cognition* 136, 150–155. doi: 10.1016/j.cognition.2014.11.003
- Muldoon, K., Simms, V., Towse, J., Menzies, V., and Yue, G. (2011). Cross-cultural comparisons of 5-year-olds' estimating and mathematical ability. *J. Cross Cult. Psychol.* 42, 669–681. doi: 10.1177/0022022111406035
- Mullis, I. V. S., Martin, M. O., Foy, P., and Hooper, M. (2016a). *TIMSS 2015 International Results in Mathematics*. Chestnut Hill, MA: TIMSS and PIRLS International Study Center, Boston College.
- Mullis, I. V. S., Martin, M. O., and Loveless, T. (2016b). *20 Years of TIMSS: International Trends in Mathematics and Science Achievement, Curriculum, and Instruction*. Chestnut Hill, MA: TIMSS and PIRLS International Study Center, Boston College.
- Nickerson, R. S. (1988). Counting, computing, and the representation of numbers. *J. Hum. Factors* 30, 181–199. doi: 10.1177/001872088803000206
- Nuerk, H.-C., Weger, U., and Willmes, K. (2005). Language effects in magnitude comparison: small, but not irrelevant. *Brain Lang.* 92, 262–277. doi: 10.1016/j.bandl.2004.06.107
- Okamoto, Y. (2015). "Mathematics learning in the USA and Japan: influences of language," in *Oxford Handbook of Mathematical Cognition*, eds R. Cohen Kadosh and A. Dowker (Oxford: Oxford University Press), 415–429.
- Perry, M., Vanderstoep, S. W. and Yu, S. L. (1993). Asking questions in first-grade mathematics classes: potential influences on mathematical thought. *J. Educ. Psychol.* 85, 31–40. doi: 10.1037/0022-0663.85.1.31
- Petitto, A. L. (1990). Development of number line and measurement concepts. *Cogn. Instr.* 7, 55–78. doi: 10.1207/s1532690xci0701_3
- Pixner, S., Zuber, J., Heřmanová, V., Kaufmann, L., Nuerk, H.-C., and Moeller, K. (2011). One language, two number-word systems and many problems: numerical cognition in the Czech language. *Res. Dev. Disabil.* 32, 2683–2689. doi: 10.1016/j.ridd.2011.06.004
- Sasanguie, D., Göbel, S. M., Moll, K., Smets, K., and Reynvoet, B. (2013). Approximate number sense, symbolic number processing, or number-space mappings: what underlies mathematics achievement? *J. Exp. Child Psychol.* 114, 418–431. doi: 10.1016/j.jecp.2012.10.012
- Saxton, M., and Towse, J. N. (1998). Linguistic relativity: the case of place value in multi-digit numbers. *J. Exp. Child Psychol.* 69, 66–79. doi: 10.1006/jecp.1998.2437
- Schneider, M., Grabner, R. H., and Paetsch, J. (2009). Mental number line, number line estimation, and mathematical achievement: their interrelations in grades 5 and 6. *J. Educ. Psychol.* 101, 359–372. doi: 10.1037/a0013840
- Schneider, M., Heine, A., Thaler, V., Torbeyns, J., De Smedt, B., Verschaffel, L., et al. (2008). A validation of eye movements as a measure of elementary school children's developing number sense. *Cogn. Dev.* 23, 409–422. doi: 10.1016/j.cogdev.2008.07.002
- Schneider, M., Merz, J., Sticker, J., De Smedt, B., Torbeyns, J., Verschaffel, L., et al. (2018). Associations of number line estimation with mathematical competence: a meta-analysis. *Child Dev.* 89, 1467–1484. doi: 10.1111/cdev.13068
- Siegler, R. S., and Booth, J. L. (2004). Development of numerical estimation in young children. *Child Dev.* 75, 428–444. doi: 10.1111/j.1467-8624.2004.00684.x
- Siegler, R. S., and Mu, Y. (2008). Chinese children excel on novel mathematical problems even before elementary school. *Psychol. Sci.* 19, 759–763. doi: 10.1111/j.1467-9280.2008.02153.x
- Simms, V., Clayton, S., Cragg, L., Gilmore, C., and Johnson, S. (2016). Explaining the relationship between number line estimation and mathematical

- achievement: the role of visuomotor integration and visuospatial skills. *J. Exp. Child Psychol.* 145, 22–33. doi: 10.1016/j.jecp.2015.12.004
- Stigler, J., Fernandez, C., and Yoshida, M. (1996). “Traditions of mathematics in Japanese and American elementary classrooms,” in *Theories of Mathematical Learning*, eds L. Steffe, P. Nesher, P. Cobb, G. Goldin, and B. Greer (Mahwah, NJ: Erlbaum), 149–175
- Sturman, L. (2015). “What is there to learn from international surveys of mathematics achievement?,” in *Oxford Handbook of Mathematical Cognition*, eds R. Cohen Kadosh and A. Dowker (Oxford: Oxford University Press), 430–444.
- Van Rinsveld, A., and Schiltz, C. (2016). Sixty-twelve = Seventy-two? A cross-linguistic comparison of children’s number decoding. *Br. J. Dev. Psychol.* 34, 181–197. doi: 10.1111/bjdp.12122
- Wong, N. Y., Lam, C. C., Wong, K. M. P., Leung, F. K. S., and Mok, I. A. C. (2001). Students’ views of mathematics learning: a cross-sectional survey in Hong Kong. *Educ. J.* 29, 38–59.
- Zhang, J., and Norman, D. A. (1995). A representational analysis of numeration systems. *Cognition* 57, 271–295. doi: 10.1016/0010-0277(95)00674-3
- Zhang, J., and Wang, H. (2005). The effect of external representations on numeric tasks. *Q. J. Exp. Psychol.* 58A, 817–838. doi: 10.1080/02724980443000340
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Dowker and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Response: Commentary: The Developmental Trajectory of the Operational Momentum Effect

Daniele Didino^{1*}, Pedro Pinheiro-Chagas², Guilherme Wood^{3,4} and André Knops^{1,5,6}

¹ Department of Psychology, Faculty of Life Sciences, Humboldt-Universität zu Berlin, Berlin, Germany, ² Laboratory of Behavioral and Cognitive Neuroscience, Stanford Human Intracranial Cognitive Electrophysiology Program, Department of Neurology and Neurological Sciences, Stanford University, Stanford, CA, United States, ³ Department of Psychology, University of Graz, Graz, Austria, ⁴ BioTechMed-Graz, University of Graz, Graz, Austria, ⁵ CNRS UMR 8240, Laboratory for the Psychology of Child Development and Education, Paris, France, ⁶ University Paris Descartes, Sorbonne Paris Cité, Paris, France

Keywords: operational momentum, mental arithmetic, numerical cognition, spatial biases, approximate mental calculation, number-to-line mapping tasks

A Commentary on

Commentary: The Developmental Trajectory of the Operational Momentum Effect

by Fischer, M. H., Miklashevsky, A. A., and Shaki, S. (2018). *Front. Psychol.* 9:2259. doi: 10.3389/fpsyg.2018.02259

OPEN ACCESS

Edited by:

Elena Nava,
Università degli Studi di Milano
Bicocca, Italy

Reviewed by:

John Opfer,
The Ohio State University,
United States

*Correspondence:

Daniele Didino
daniele.didino@gmail.com

Specialty section:

This article was submitted to
Developmental Psychology,
a section of the journal
Frontiers in Psychology

Received: 07 December 2018

Accepted: 16 January 2019

Published: 06 February 2019

Citation:

Didino D, Pinheiro-Chagas P, Wood G
and Knops A (2019) Response:
Commentary: The Developmental
Trajectory of the Operational
Momentum Effect.
Front. Psychol. 10:160.
doi: 10.3389/fpsyg.2019.00160

Fischer et al. (2018) (henceforth: FM&S) raised theoretical and methodological criticisms against our study (Pinheiro-Chagas et al., 2018) on the development of the operational momentum effect (OM). Here, we will refute their criticisms and argue for a more precise definition of the OM as the operation-induced misestimation of arithmetic problem outcomes.

First, FM&S advocate the idea that zero-problems (e.g., 6+0) would be ideally suited to reveal OM. FM&S ask “how does [the attentional shift] account explain larger OM with zero problems?” In Pinhas and Fischer (2008) task, zero problems only required to map a number (the first operand) onto a labeled line, since these problems are solved by means of rules (i.e., $N+0 = N$, $N-0 = N$) rather than mental calculation (Butterworth et al., 2001; Campbell and Metcalfe, 2007). Therefore, FM&S’s question is not valid because its premise (i.e., zero-problems produce OM) is not valid. Since zero and non-zero problems do not invoke the same strategies, merging their respective biases will not be helpful in elucidating the underlying mechanisms. The attentional shift account aims to describe the operation-specific outcome misestimations caused by mentally combining (at least) two numerosities. FM&S further argue that a stronger bias for zero problems compared to non-zero problems (Pinhas and Fischer, 2008; Shaki et al., 2018) invalidates the compression account of the OM “because the logarithm of zero is not defined.” This argumentation is flawed because FM&S mix up logarithm as a mathematical function (not defined for zero, indeed) with logarithm as a model (coding scheme) to describe the compressed internal scale of the representation of magnitudes (Nieder and Miller, 2003; Harvey et al., 2013). In the latter case, the logarithmic function is used as mathematical approximation of the relation between external physical magnitude and its internal representation. However, it makes no sense to assume that cortical circuits actually compute the faithful “mathematical log transformation” of a given sensory information. The intensity of external physical stimuli is internally represented via non-linear spatio-temporal neural codes (e.g., rate code, population code). Basing their criticism on the restriction of the mathematical definition of the logarithm to positive real numbers, FM&S conflate the mathematical definition with the neural and cognitive representation of magnitudes. Moreover, even assuming that the cognitive

system would actually be bound to this particular *mathematical* formulation of the relation between physical stimulus magnitude and sensation, another framework has been put forward that does define a mathematical solution of zero magnitudes. Stevens's power function (with positive real exponents smaller than 1) can provide identical predictions and is defined for zero. In sum, the fact that "*the logarithm of zero is not defined*" does not invalidate the compression account nor seems the use of zero problems ideal for investigating OM.

Second, we argued that the attentional shift account and the heuristic account provide equivalent predictions. Fischer and colleagues criticize this by stating that it is in conflict with results from McCrink and Hubbard and cite: "... *the use of heuristics is generally increased when attention is decreased*" (McCrink and Hubbard, 2017, p. 240). Our interpretation of McCrink and Hubbard's manuscript was based on the idea that these two accounts "*are actually so deeply intertwined that they are indistinguishable*" (p. 240) and on the fact that McCrink and Hubbard's findings "*can be best described with a heuristics-via-spatial-shifts account*" (p. 241).

Third, FM&S criticize that the downward (upward) movement of addends (subtrahends) would be inconsistent with "*the vertical MNL*" and ask "*why [...] operations along a horizontal MNL [were] primed with vertical movements?*" We argue that these movements actually mimic our daily experience: adding objects from the top into a box (downward movement) and subtracting them from inside a box to the top (upward movements). Any effect of this supposed inconsistency between physical vertical movements of the operands and attentional movement on the MNL should have weakened, eliminated or even reversed the OM. Yet, we did not observe such interference. They also reasoned that the center-to-top movement of the subtrahends "*removed attention from the place of mentally simulating the outcome, thus impeding subtraction.*" First, this conclusion is inconsistent with findings from previous studies (McCrink et al., 2007; McCrink and Hubbard, 2017), where OM was observed despite subtrahends moving to the right (i.e., inconsistently with the *horizontal MNL*). Second, FM&S conflate mental simulation of addition and subtraction with attentional focus in external space. After all, the outcomes are estimated

in the participants' minds—not in external space where no numerical information is present at that point in time.

Finally, the idea that in our previous studies "*the normal ingredients of OM are dis-ordered or diluted*" originates from the divergent definition of the OM. In line with the original definition by McCrink et al. (2007), we propose that OM emerges during mental calculation, rather than rule application or arithmetic fact retrieval, and refers to the numerical deviation in estimated outcomes of arithmetic operations (e.g., addition vs. subtraction), rather than biases resulting from mapping outcomes to a non-numerical dimension. In number-to-line mapping tasks, participants locate addition and subtraction outcomes on a labeled line (Pinhas and Fischer, 2008) or modify the length of a line proportionally to addition and subtraction outcomes (Shaki et al., 2015, 2018). These paradigms do not measure outcome deviations, but rather they require an additional transformation process where the outcome is converted into another physical dimension (number to position or length). Both tasks can be subject to strategic (e.g., use of reference points; Barth and Paladino, 2011; Slusser et al., 2013; Sasanguie et al., 2016; but see Opfer et al., 2016) or procedural biases (e.g., perceptual hysteresis). Therefore, any observed biases may arise from the additional transformation process rather than the calculation process itself. Results from procedures that analyse only the final location on a labeled line (Pinhas and Fischer, 2008) or the length of a segment (Shaki et al., 2015, 2018) must be interpreted cautiously because they are not measuring OM but biases that may well take place after the calculation process and have their origin in the transformation algorithm.

AUTHOR CONTRIBUTIONS

DD and AK wrote the manuscript. PP-C and GW provided critical revision. All authors approved the final version of the manuscript for submission.

ACKNOWLEDGMENTS

This work was supported by a grant (DI 2361/1-1) from Deutsche Forschungsgemeinschaft (DFG, German Research Council) to DD.

REFERENCES

- Barth, H. C., and Paladino, A. M. (2011). The development of numerical estimation: evidence against a representational shift. *Dev. Sci.* 14, 125–135. doi: 10.1111/j.1467-7687.2010.00962.x
- Butterworth, B., Zorzi, M., Girelli, L., and Jonckheere, A. R. (2001). Storage and retrieval of addition facts: the role of number comparison. *Q. J. Exp. Psychol. Sect. A* 54, 1005–1029. doi: 10.1080/713756007
- Campbell, J. I. D., and Metcalfe, A. W. S. (2007). Arithmetic rules and numeral format. *Eur. J. Cogn. Psychol.* 19, 335–355. doi: 10.1080/09541440600717610
- Fischer, M. H., Miklashevsky, A., and Shaki, S. (2018). Commentary: the developmental trajectory of the operational momentum effect. *Front. Psychol.* 9:2259. doi: 10.3389/fpsyg.2018.02259
- Harvey, B. M., Klein, B. P., Petridou, N., and Dumoulin, S. O. (2013). Topographic representation of numerosity in the human parietal cortex. *Science* 341, 1123–1126. doi: 10.1126/science.1239052
- McCrink, K., Dehaene, S., and Dehaene-Lambertz, G. (2007). Moving along the number line: operational momentum in nonsymbolic arithmetic. *Percept. Psychophys.* 69, 1324–1333. doi: 10.3758/BF03192949
- McCrink, K., and Hubbard, T. (2017). Dividing attention increases operational momentum. *J. Numer. Cogn.* 3, 230–245. doi: 10.5964/jnc.v3i2.34
- Nieder, A., and Miller, E. K. (2003). Coding of cognitive magnitude: compressed scaling of numerical information in the primate prefrontal cortex. *Neuron* 37, 149–157. doi: 10.1016/S0896-6273(02)01144-3
- Opfer, J. E., Thompson, C. A., and Kim, D. (2016). Free versus anchored numerical estimation: a unified approach. *Cognition* 149, 11–17. doi: 10.1016/j.cognition.2015.11.015
- Pinhas, M., and Fischer, M. H. (2008). Mental movements without magnitude? A study of spatial biases in symbolic arithmetic. *Cognition* 109, 408–415. doi: 10.1016/j.cognition.2008.09.003

- Pinheiro-Chagas, P., Didino, D., Haase, V. G., Wood, G., and Knops, A. (2018). The developmental trajectory of the operational momentum effect. *Front. Psychol.* 9:1062. doi: 10.3389/fpsyg.2018.01062
- Sasanguie, D., Verschaffel, L., Reynvoet, B., and Luwel, K. (2016). The development of symbolic and non-symbolic number line estimations: three developmental accounts contrasted within cross-sectional and longitudinal data. *Psychol. Belg.* 56, 382–405. doi: 10.5334/pb.276
- Shaki, S., Pinhas, M., and Fischer, M. H. (2018). Heuristics and biases in mental arithmetic: revisiting and reversing operational momentum. *Think. Reason.* 24, 138–156. doi: 10.1080/13546783.2017.1348987
- Shaki, S., Sery, N., and Fischer, M. H. (2015). $1 + 2$ is more than $2 + 1$: violations of commutativity and identity axioms in mental arithmetic. *J. Cogn. Psychol.* 27, 471–477. doi: 10.1080/20445911.2014.973414
- Slusser, E. B., Santiago, R. T., and Barth, H. C. (2013). Developmental change in numerical estimation. *J. Exp. Psychol. Gen.* 142, 193–208. doi: 10.1037/a0028560
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Didino, Pinheiro-Chagas, Wood and Knops. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers



OPEN ACCESS

Articles are free to read
for greatest visibility
and readership



FAST PUBLICATION

Around 90 days
from submission
to decision



HIGH QUALITY PEER-REVIEW

Rigorous, collaborative,
and constructive
peer-review



TRANSPARENT PEER-REVIEW

Editors and reviewers
acknowledged by name
on published articles

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org

Contact us: info@frontiersin.org | +41 21 510 17 00



REPRODUCIBILITY OF RESEARCH

Support open data
and methods to enhance
research reproducibility



DIGITAL PUBLISHING

Articles designed
for optimal readership
across devices



FOLLOW US

@frontiersin



IMPACT METRICS

Advanced article metrics
track visibility across
digital media



EXTENSIVE PROMOTION

Marketing
and promotion
of impactful research



LOOP RESEARCH NETWORK

Our network
increases your
article's readership