



Can I Influence You? Development of a Scale to Measure Perceived Persuasiveness and Two Studies Showing the Use of the Scale

Rosemary J. Thomas^{1,2*}, Judith Masthoff^{1,3*} and Nir Oren^{1*}

¹ Computing Science, University of Aberdeen, Aberdeen, United Kingdom, ² Institute of Global Food Security, Queen's University Belfast, Belfast, United Kingdom, ³ Computing Science, Utrecht University, Utrecht, Netherlands

OPEN ACCESS

Edited by:

Ralf Klamma,
RWTH Aachen University, Germany

Reviewed by:

Jorge Luis Bacca Acosta,
Fundación Universitaria Konrad
Lorenz, Colombia
Kirsten Ailsa Smith,
University of Southampton,
United Kingdom

*Correspondence:

Rosemary J. Thomas
rosemaryjthomas@acm.org
Judith Masthoff
j.f.m.masthoff@uu.nl
Nir Oren
n.oren@abdn.ac.uk

Specialty section:

This article was submitted to
AI for Human Learning and Behavior
Change,
a section of the journal
Frontiers in Artificial Intelligence

Received: 21 May 2019

Accepted: 25 October 2019

Published: 21 November 2019

Citation:

Thomas RJ, Masthoff J and Oren N
(2019) Can I Influence You?
Development of a Scale to Measure
Perceived Persuasiveness and Two
Studies Showing the Use of the Scale.
Front. Artif. Intell. 2:24.
doi: 10.3389/frai.2019.00024

In this paper, we develop and validate a scale to measure the perceived persuasiveness of messages to be used in digital behavior interventions. A literature review is conducted to inspire the initial scale items. The scale is developed using Exploratory and Confirmatory Factor Analysis on the data from a study with 249 ratings of healthy eating messages. The construct validity of the scale is established using ratings of 573 email security messages. Using the data from the two studies, we also show the usefulness of the scale by analyzing the perceived persuasiveness of different message types on the developed scale factors in both the healthy eating and email security domains. The results of our studies also show that the persuasiveness of message types is domain dependent and that when studying the persuasiveness of message types, the finer-grained argumentation schemes need to be considered and not just Cialdini's principles.

Keywords: perceived persuasiveness, scale development, behavior change, message type, argumentation schemes

1. INTRODUCTION

Many behavior change interventions have been developed for a wide variety of domains. For example, "Fit4Life" (Purpura et al., 2011) promotes healthy weight management, the ASICA application (Smith et al., 2016) reminds skin-cancer patients to self-examine their skin, the SUPERHUB application (Wells et al., 2014) motivates sustainable travel, while "Portia" (Mazzotta et al., 2007) and "Daphne" (Grasso et al., 2000) encourage healthy eating habits.

Clearly, it is important to measure the effectiveness of such persuasive interventions. However, it is often difficult to measure actual persuasiveness (O'Keefe, 2018). Perhaps the primary three reasons for such difficulties are as follows. First, measuring actual persuasiveness tends to require more time and effort from participants and additional resources. For example, to measure the persuasiveness of a healthy eating intervention, participants may need to provide detailed diaries of their food intake, which are cumbersome and often unreliable (Cook et al., 2000), and may require the provisioning of scales to participants. Also, when studying many experimental conditions, it may be hard to obtain sufficient participants willing to spend the necessary time [e.g., to measure actual persuasiveness of reminders in (Smith et al., 2016) would have required a large number of skin cancer patients]. Second, it is hard to measure actual persuasiveness due to confounding factors. For example, when measuring the persuasiveness of a sustainable transport application,

other factors such as the weather may influence people's behavior. Third, there may be ethical issues which make it hard to measure actual persuasiveness. For example, if one wanted to investigate the persuasive effects of different message types to get learners to study more, it may be deemed unethical to do this in a real class room, as learners in the control condition may be seen to be disadvantaged. Purpura et al. (2011) illustrates some of the ethical problems while using persuasive technologies in behavior change interventions.

Because of these difficulties in measuring actual persuasiveness, *perceived* persuasiveness is often used as an approximation of, or the initial step in the measurement of, actual persuasiveness (see **Table 1** for example studies that used perceived persuasiveness). Perceived persuasiveness may include multiple factors. For example, perceived effectiveness in changing somebody's attitudes may be different from perceived effectiveness in changing behavior. We would like a reliable scale that incorporates multiple factors as sub-scales, with each sub-scale consisting of multiple items. Such a scale does not yet exist, and researchers have so far had to use their own measures without proper validation.

Therefore, this paper describes the process for developing a reliable and validated multi-item, multi-subscale scale to measure perceived persuasiveness. In addition, the data collected will be used to show the usefulness of the scale by analyzing the impact of different persuasive message types on the developed scale factors.

2. LITERATURE REVIEW

To inspire the scale items and show the need for scale development, we first investigated how researchers measured perceived persuasiveness by examining the scale items and respective measurements they used in published user studies. We performed a semi-structured literature review, searching in Scopus from the period 2014 to 2018 across disciplines. At first we performed a narrow search using the following search query:

"scales development" AND studies AND persuasion.

However, this produced very few search results. Later, we modified the search query to the following:

persuasion AND (experiments OR studies)

to get a broader range of articles. We also searched in the Proceedings of the "International Conference on Persuasive Technology" for the period from 2013 to 2018. We were looking for user studies that developed or used a scale to measure perceived persuasiveness. The search resulted in 12 papers, including 2 from outside computer science from marketing and communications (Koch and Zerbac, 2013; Zhang et al., 2014). Ham et al. (2015) and O'Keefe (2018) appeared in the initial search results but were excluded as they contained meta-reviews rather than original studies. Three papers were added to the results through snowballing, given these specifically addressed perceived persuasiveness scales:

- Kaptein et al. (2009) cited in Busch et al. (2013).
- MacKenzie and Lutz (1989) cited in Ham et al. (2015).
- Zhao et al. (2011) cited in O'Keefe (2018).

The results of the literature search are shown in **Table 1**, which lists 60 scale items and their measurements based on studies reported in these 15 papers¹.

Unfortunately, most studies do not report on the scale construction, reliability or validation. The exceptions are Kaptein et al. (2009) and Busch et al. (2013). However, Kaptein et al. (2009)'s scale really measures the susceptibility of participants to certain Cialdini's principles of persuasion (such as liking and authority) (Cialdini, 2009), rather than the persuasion of the messages themselves. Similarly, Busch et al. (2013) aims to measure the persuasibility of participants by certain persuasive strategies (such as social comparison and rewards).

We reduced the 60 items listed in **Table 1** in two steps. First, we removed duplicates and merged highly similar items. Next, we transformed items that were not yet related to a message where possible (items 9, 11–13, 35–36). For instance, item 11 "This feature would make me more aware of [policy]" was changed into "This message makes me more aware of my behavior," and item 35 "I always follow advice from my general practitioner" was changed into "I will follow this message." Finally, we removed items for which this was not possible (e.g., items 37–44 that measure a person's susceptibility, and items such as 10, 55). This reduced the list to the 30 items used for the initial scale development as shown in **Table 2**, which also shows which original items these were derived from.

A limitation of our systematic literature review is that it was mainly restricted to papers published in the period 2014–2018². Additionally, it is possible for a systematic review to miss papers due to the search terms used or the limitation of searching abstracts, titles, and keywords. Some other papers related to measuring persuasiveness were found after the review was completed, most noticeably (Feltham, 1994; Allen et al., 2000; Lehto et al., 2012; Popova et al., 2014; Jasek et al., 2015; Yzer et al., 2015; McLean et al., 2016). We will discuss how the scales developed in this paper relate to this other work in our discussion section.

3. STUDY DESIGN

3.1. Study 1: Development of a Perceived Persuasiveness Scale

We conducted a study to develop a rating scale to measure the "perceived persuasiveness" of messages. The aim was to obtain a scale with good internal consistency, and with at least three items per factor following the advice in MacCallum et al. (1999) to have at least three or four items with high loadings per factor.

¹Many of these papers contained additional items; these were normally not related to measuring persuasiveness.

²In this period much research on persuasive technology has taken place, as evidenced by 7,410 papers being found for "persuasive technology" in Google Scholar.

TABLE 1 | Scale items related to measuring perceived persuasiveness, the measurement scale used for each item, and the number of measurement points.

References	Number	Scale items	Scale measurement	Points
Anagnostopoulou et al. (2017)	1 2 3 4	The [System] would Influence me be convincing be personally relevant for me make me [target behavior]	Strongly disagree to strongly agree	7
Thomas et al. (2017)	5 6 7 8	Motivational Appropriateness Effectiveness Convincing	Not very motivating to very motivating Very inappropriate to very appropriate Very ineffective to very effective Very unconvincing to very convincing	5
Busch et al. (2016)	9 10 11 12 13	I find this feature useful I enjoy using this feature This feature would make me more aware of [policy] have a positive influence on my attitude toward [policy] lead me to comply with [policy].	Strongly disagree to strongly agree	7
Oduor and Oinas-Kukkonen (2017)	14 15 16 17	The system provides trustworthy content believable content accurate content professional information	Strongly disagree to strongly agree	7
Chang et al. (2018)	18 19 20 21	On average, [communications] are persuasive compelling logical plausible	Strongly disagree to strongly agree	5
Zhao et al. (2011)	22 23 24	[Communication] that is believable convincing important to me	Strongly disagree to strongly agree	5
Orji (2014), Orji et al. (2014)	25 26 27 28	The system would influence me be convincing be personally relevant for me make me reconsider my [behavior]	Strongly disagree to strongly agree	7
Zhang et al. (2014)	29 30 31 32 33 34	[Communications] were convincing persuasive strong good trustworthy reliable	Strongly disagree to strongly agree	7
Kaptein et al. (2009) ^a	35 36 37	Susceptibility authority: I always follow advice from my general practitioner. When a professor tells me something I tend to believe it is true. Susceptibility consensus: If someone from my social network notifies me about a good book, I tend to read it.	Totally disagree to Totally agree	7

(Continued)

TABLE 1 | Continued

References	Number	Scale items	Scale measurement	Points
	38	When I am in a new situation I look at others to see what I should do.		
		Susceptibility liking:		
	39	I accept advice from my social network.		
	40	When I like someone, I am more inclined to believe him or her.		
Busch et al. (2013)	41	Before I do something, I want to know how other people have done it, so I can feel more safe.	Fully agree to Fully disagree	9
	42	It is important to me to know what other people are doing.		
	43	I trust information better where the source is specified.		
	44	It is important for me to be precisely informed about things that I need to do, before I do them.		
Hammer et al. (2016)	45	[Communications were]	Not polite to Very polite Not persuasive at all to Very persuasive	7
Hossain and Saini (2014)	46	The [communication] is	Truthful to Not truthful Unbelievable to Believable Not deceptive to Deceptive	8
	47	The [communicator] is	Sincere to Insincere Honest to Dishonest Not manipulative to Manipulative Not pushy to Pushy	
Meschtscherjakov et al. (2016)		This system	Completely disagree to completely agree	7
	48	makes people change their behavior		
	49	has the potential to influence people		
	50	gives the behavior of its users a new direction		
	51	is exactly what I need to change my attitude		
	52	does not cause a change in behavior with me		
	53	causes me to do some things differently		
	54	Thanks to the system I reach my goals.		
	55	I will use this system as often as possible.		
	56	With the help of the system, I will behave differently in the future.		
Koch and Zerbac (2013)	57	I had the feeling that [communicator] wanted to convince the reader of [communicator]'s standpoint	I do not agree at all to I fully agree	5
	58	[Communicator] wanted to convince me of [communicator]'s views		
MacKenzie and Lutz (1989)	59	Attitude: The [communicator/communication] is	Good to bad Pleasant to Unpleasant Favorable to Unfavorable	7
	60	Credibility: The [communicator/communication] is	Convincing to Unconvincing Believable to Unbelievable Biased to Unbiased	

^aThis is a sample. They also present items related to the susceptibility to the other Cialdini principles.

3.1.1. Participants

The participants for this study were recruited by sharing the link of the study via social media and mailing lists. The study had four validation questions to check if participants were randomly rating the scales. After removing such participants, a total of 92 participants rated 249 messages.

3.1.2. Procedure

Each participant was shown a set of five messages (see Table 4), each promoting healthy eating. These messages were

based on different *argumentation schemes*³ (Walton et al., 2008) and were produced in another study using a message generation system (Thomas et al., 2018). Each message was rated using 34 scale items (the scale items marked with * act as validation checks) on a 7-point Likert scale that ranges from “strongly disagree” to “strongly agree” (see Table 2 and Figure 1). Finally, participants were given the option to provide feedback.

³Argumentation schemes are stereotypical patterns of reasoning.

TABLE 2 | Scale items developed used in Study 1.

Scale items	Inspired from / Similar to
This message is	
influencing.®	1,25
convincing. ®	2,8,23,26,29,57,58
personally relevant.®	3,24,27
motivating. ®	5
appropriate.	6
credible.	21,31,60
encouraging. ®	7,18,30,45
inappropriate.*	N/A
effective.®	7
useful.®	9,17,32
believable. ®	15,22,34,46,47
ineffective.*	N/A
accurate.	16,17,20
trustworthy.	14,33,46,47
exactly what I need to help reach my goals.	54
exactly what I need to change my attitude.	49,51
exactly what I need to change my behavior.	51
This message	
makes me more aware of my behavior.®	11
leads me to comply with behavior expectations.®	13,19
will cause changes in my behavior.	4
has a positive influence on my attitude.®	12
has the potential to change user behavior.	48
has the potential to influence user behavior.	49
has the potential to inspire users.	50
causes a change in my behavior.	4,52
causes me to make some changes in my behavior.	4,53
I	
will follow this message.®	13,19,35
consider this message.	28
accept this message.	39
believe this message is true.®	36
After viewing this message, I will make	
some behavior change in the future.	56
changes in my attitude.	49,51
Please click the second option from the	
right.*	N/A
left.*	N/A

*act as validation check. ®cross loaded on different factors.

3.1.3. Research Question and Hypothesis

We were interested in the following research question:

- RQ1: What is a reliable scale to measure perceived persuasiveness?

In addition, we wanted to investigate the usefulness of the scale by analyzing whether the different message types had an impact on the ratings of the developed factors. Therefore, we formulated the following hypothesis:

TABLE 3 | Study 1: Reduced scales items after EFA.

Factors	Scale items
Effectiveness	This message will cause changes in my behavior.® This message is exactly what I need to help reach my goals.® This message is exactly what I need to change my attitude.® This message is exactly what I need to change my behavior.® This message causes a change in my behavior. This message causes me to make some changes in my behavior. After viewing this message, I will make some behavior change in the future.® After viewing this message, I will make changes in my attitude.
Quality	This message is appropriate.® This message is credible.® This message is believable. This message is accurate. This message is trustworthy. I believe this message is true. I accept this message.®
Capability	This message has the potential to change user behavior. This message has the potential to influence user behavior. This message has the potential to inspire users.

® high Standardized Residual Covariances with several other items.

TABLE 4 | Healthy eating messages used in Study 1 with corresponding argumentation schemes.

Scheme name	Message
Argument from commitment with goal	As you want to eat healthy, you are committed to eating healthy foods. So, you are also committed to shopping carefully and reading the labels as it helps you to eat healthy foods.
Argument from expert opinion with goal	A nutritionist recommends that you keep a log of your daily calorie intake to manage calorie intake. So, you should follow their recommendation.
Argument from position to know with goal	A college football star suggests that you eat a diet high in protein to have more energy. So, you should follow their suggestion.
Argument from sunk cost with action	You have a choice whether or not to eat vegetables with every serving, however, you committed to doing so earlier. So, you should choose to eat vegetables with every serving.
Practical reasoning with goal	If you cut out added sugars, white flours, white rice and soft drinks, it helps you to lose weight. So, you ought to do this.

- H1: Perceived persuasiveness of each factor differs for different message types.

3.2. Study 2: Validation of the Perceived Persuasiveness Scale

Next, we conducted a study to determine the construct validity of the developed scale. We replicated the scale-testing in the domain of email security using another data set.

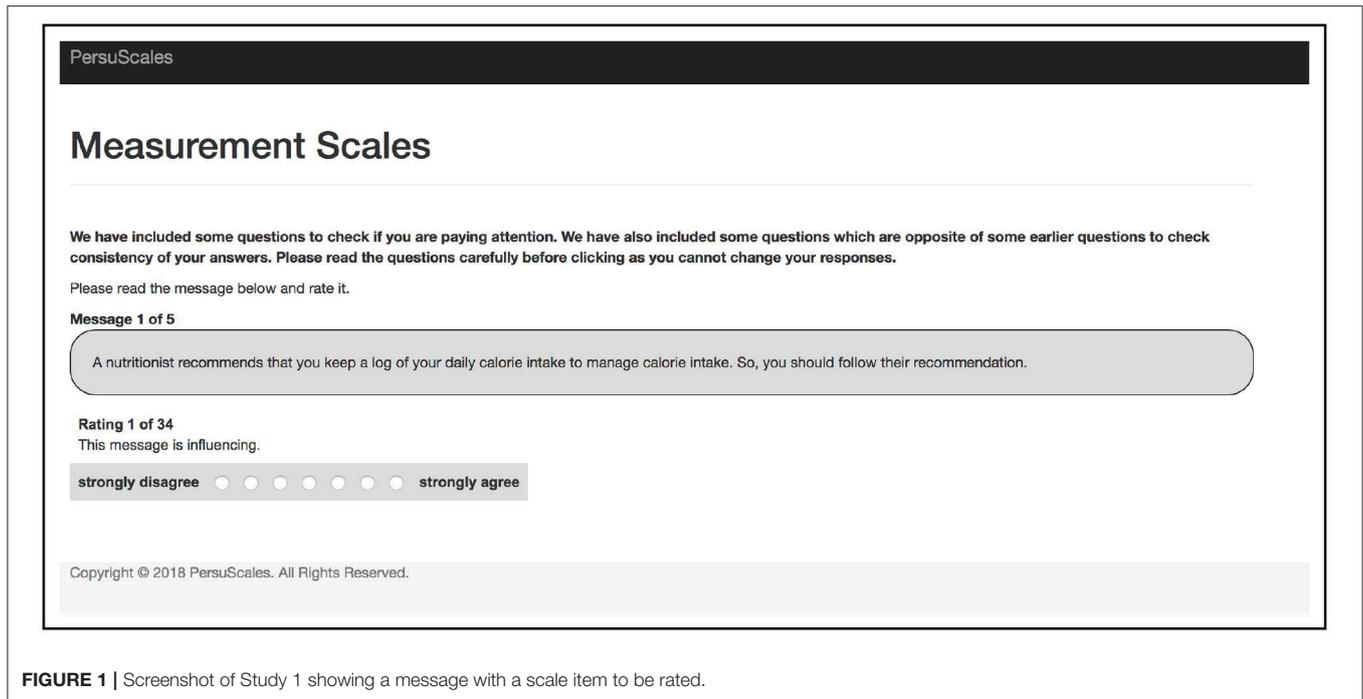


FIGURE 1 | Screenshot of Study 1 showing a message with a scale item to be rated.

TABLE 5 | Email security messages used in Study 2 with corresponding argumentation schemes.

Scheme name	Message
Argument from commitment with goal	As you want to keep your computer account safe, you are committed to check whether website links are genuine in emails. So, you are also committed to preview website links in an email application as it helps you to check whether website links are genuine in emails.
Argument from expert opinion with goal	A renowned email security expert recommends that you prevent opening suspicious attachments to protect your email account. So, you should follow their recommendation.
Argument from position to know with goal	A colleague who attended email security training suggests that you verify the logo, header and footer of email newsletters to make sure that they originate from genuine sources. So, you should follow their suggestion.
Argument from sunk cost with action	You have a choice whether or not to be security-conscious when processing email; however, you committed to doing so earlier. So, you should choose to be security-conscious when processing email.
Practical reasoning with goal	If you choose not to provide personal information by responding to emails that threaten to disable account access, it helps you to safeguard your email account. So, you ought to do this.

3.2.1. Participants

The participants for this study were recruited by sharing the link of the study via social media and mailing lists. After removing the invalid participants (as before), a total of 134 participants rated 573 messages.

TABLE 6 | Study 1: Reduced scale items after CFA.

Factors	Scale items
Effectiveness	This message will cause changes in my behavior. This message causes me to make some changes in my behavior. After viewing this message, I will make changes in my attitude.
Quality	This message is accurate. This message is trustworthy. I believe this message is true.
Capability	This message has the potential to change user behavior. This message has the potential to influence user behavior. This message has the potential to inspire users.

3.2.2. Procedure

Each participant was shown a set of five messages (see **Table 5**) that promote email security, again based on argumentation-schemes. Each message was rated using the scale (see **Table 6** and **Figure 2**) that resulted from Study 1. Finally, participants were given the option to provide feedback.

3.2.3. Research Question and Hypotheses

We were interested in the following research question:

- RQ2: How valid is the developed perceived persuasiveness scale?

Our first study: Development of a Perceived Persuasiveness Scale resulted in a scale with three factors for measuring

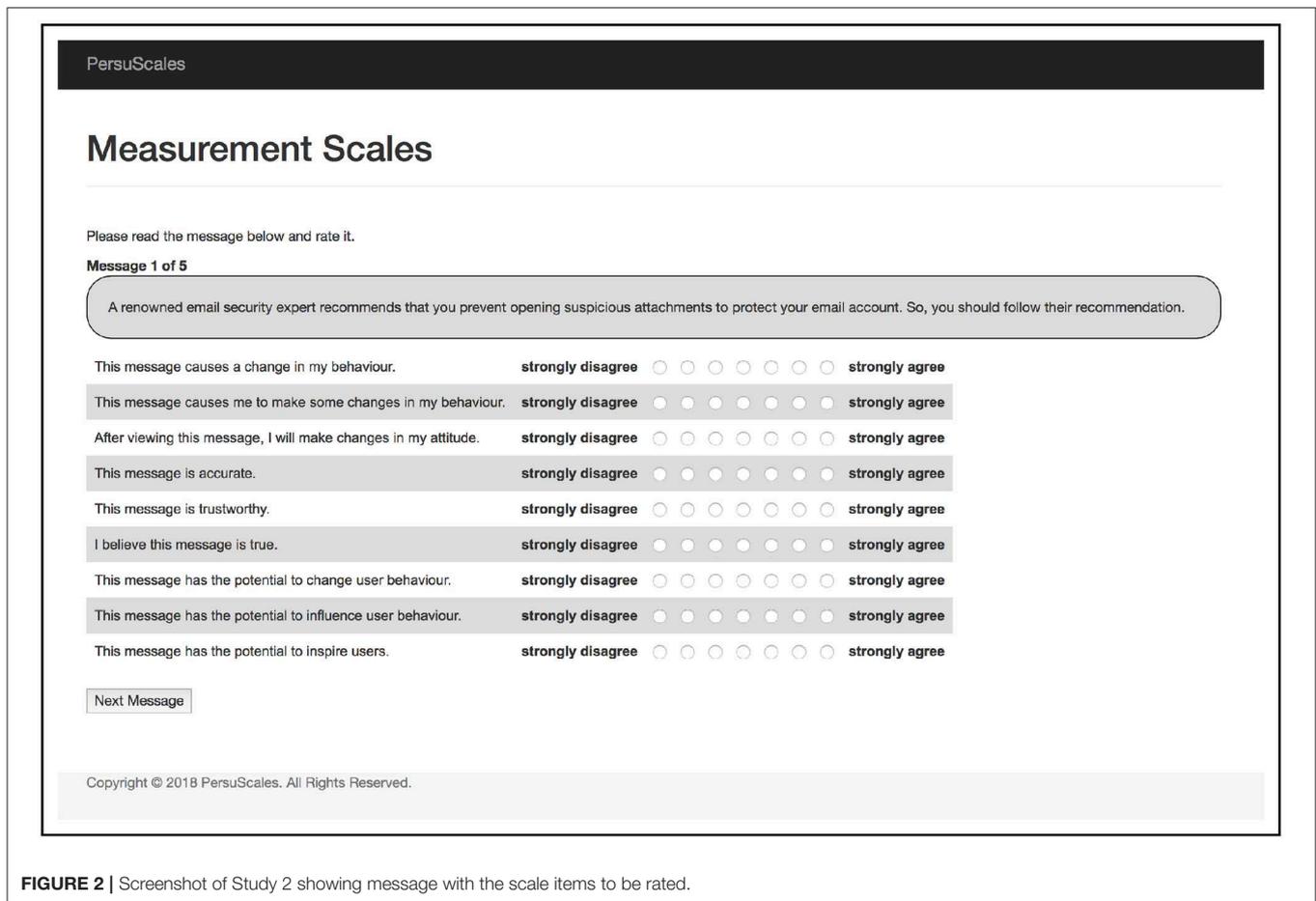


FIGURE 2 | Screenshot of Study 2 showing message with the scale items to be rated.

perceived persuasiveness: Effectiveness, Quality, and Capability (see section 4.1). We wanted to investigate the usefulness of this scale by analyzing whether the message types differed on these three developed factors. Therefore, we formulated the following hypotheses:

- H2: The perceived persuasiveness factor Effectiveness differs for different message types.
- H3: The perceived persuasiveness factor Quality differs for different message types.
- H4: The perceived persuasiveness factor Capability differs for different message types.
- H5: Overall perceived persuasiveness⁴ differs for different message types.

4. RESULTS

4.1. Study 1: Development of a Perceived Persuasiveness Scale

First we checked the Kaiser-Meyer-Olkin Measure of Sampling Adequacy, which was greater than 0.90. According to this

⁴Overall perceived persuasiveness was calculated as the mean of the factors: Effectiveness, Quality, and Capability.

measure, values in the 0.90's indicate that the sampling adequacy is "marvelous" (Dziuban and Shirkey, 1980). Next, we investigated the inter-item correlations. For the factor analysis, all the 7-point scale items were considered as ordinal measures. To further filter the items and identify the factors, we conducted an Exploratory Factor Analysis (EFA) using Principal Component Analysis extraction and Varimax rotation with Kaiser Normalization (Howitt and Cramer, 2014). Varimax rotation was used as the matrix was confirmed orthogonal (the Component Correlation Matrix shows that the majority of the correlations was less than 0.5). We obtained three factors (see **Table 2**). The first factor we named Effectiveness as its items relate to user behavior and attitude changes and attainment of user goals. The second we named Quality as its items relate to characteristics of a message strength such as trustworthiness and appropriateness. The third we named Capability as its items relate to the *potential* for motivating users to change behavior. We removed the 13 items that cross loaded on different factors (see **Table 2** with scale items marked ®). This resulted in **Table 3**, which shows the reduced scale items for the three factors. We checked the Cronbach's Alpha of all the items belonging to the three factors separately. It was greater than 0.9 for each of the three factors which indicates "excellent" scale reliability.

Next, we conducted Confirmatory Factor Analyses (CFA) to determine the validity of the scale, and to confirm the factors and items by checking the model fit (Hu and Bentler, 1999). Based on these analyses, 8 items were removed due to high Standardized Residual Covariances with several other items which were greater than 0.4. The items removed are the items in **Table 3** marked ®.

Table 6 shows the resulting scale of 9 items. The final Confirmatory Factor Analysis resulted in the following values for the Tucker-Lewis Index (TLI) = 0.988, Comparative Fit Index (CFI) = 0.993, and Root Mean Square Error of Approximation (RMSEA) = 0.054, when extracting the three factors and their items. A cut off value nearing 0.95 for TLI and CFI (the higher the better) and a cut off value nearing 0.60 for RMSEA (the lower the better) are required to establish that there is an acceptable model fit between the hypothesized model and the observed data (Hu and Bentler, 1999; Schreiber et al., 2006). In the resulting scale, the TLI and CFI are above 0.95 and RMSEA is below 0.60, which shows an acceptable model fit. This answers research question RQ1.

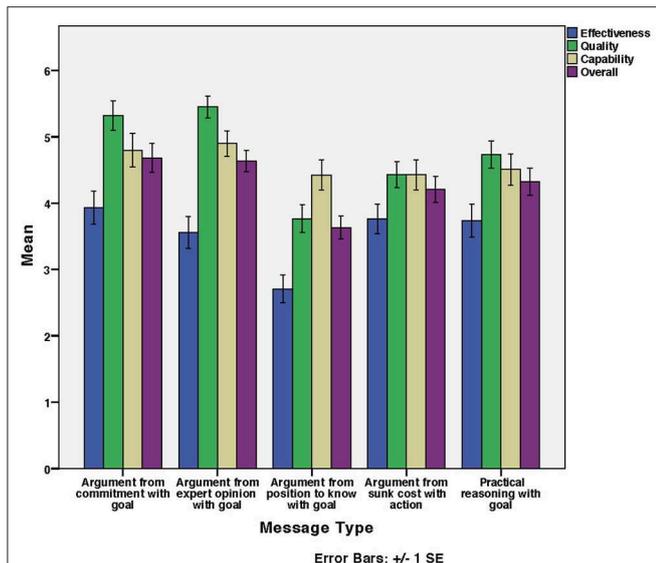


FIGURE 3 | Healthy eating messages: Mean of factors' and overall ratings for developed scale per message type.

4.2. Study 1: Impact of Message Types on Factors

Figure 3 shows the mean Effectiveness, Quality, Capability, and Overall perceived persuasiveness of message types used for the healthy eating messages. Overall perceived persuasiveness was calculated as the mean of the factors: Effectiveness, Quality, and Capability.

A one-way repeated measures MANOVA with Effectiveness, Quality, Capability, and Overall perceived persuasiveness as dependent variables and message type as the independent variable provided the results for the analyses given below. To determine the homogeneous subsets, the Ryan-Einot-Gabriel-Welch Range was selected as a *post-hoc* test since we have more than 3 levels within the independent variable (i.e., the message type).

According to Thomas et al. (2018), the argumentation schemes can be mapped to Cialdini's principles of persuasion.

1. Cialdini's Principle: Commitments and Consistency
 - Argument from commitment with goal
 - Practical reasoning with goal.
 - Argument from sunk cost with action
2. Cialdini's Principle: Authority
 - Argument from expert opinion with goal
 - Argument from position to know with goal.

The study conducted by Thomas et al. (2017) states that Authority was significantly more persuasive, followed by Commitments and Consistency and the other Cialdini principles. We were interested to know whether our findings would be similar. Hence, the analysis will consider both the argumentation schemes and Cialdini's principles when discussing the findings.

4.2.1. Impact of Message Types on Effectiveness

According to **Figure 3**, ARGUMENT FROM COMMITMENT WITH GOAL was the highest rated in Effectiveness while ARGUMENT FROM POSITION TO KNOW WITH GOAL was the lowest. There was a significant effect of message type on Effectiveness [$F(4, 244) = 4.39, p < 0.01$]. There was a significant difference between ARGUMENT FROM POSITION TO KNOW WITH GOAL and the other message types ($p < 0.05$). The rest were non-significant. **Table 7** shows the homogeneous subsets. This partially supports the hypothesis (H1) that perceived persuasiveness on each factor differs for different message types.

TABLE 7 | Study 1: Homogeneous subsets for Effectiveness, Quality, and Capability.

Message type	Effectiveness			Quality					Capability		
	Mean			Mean					Mean		
	N	S1	S2	N	S1	S2	S3	S4	N	S1	
Argument from position to know with goal	52	2.71		52	3.76					52	4.42
Argument from expert opinion with goal	52		3.56	51		4.43				51	4.42
Practical reasoning with goal	48		3.74	48		4.73	4.73			48	4.51
Argument from sunk cost with action	51		3.76	46			5.32	5.32		46	4.80
Argument from commitment with goal	46		3.93	52				5.45		52	4.90

TABLE 8 | Study 1: Homogeneous subsets for Overall Perceived Persuasiveness.

Message type	N	Mean	
		Subset 1	Subset 2
Argument from position to know with goal	52	3.63	
Argument from sunk cost with action	51	4.21	4.21
Practical reasoning with goal	48	4.32	4.32
Argument from expert opinion with goal	52		4.63
Argument from commitment with goal	46		4.68

As shown, the two Authority messages had the lowest Effectiveness scores, though the ARGUMENT FROM EXPERT OPINION WITH GOAL was not rated significantly lower than the Commitments and Consistency messages. We observe that the Effectiveness of all messages was low, below or around the mid-point of the scale. This contradicts the results from Thomas et al. (2017) where Authority and Commitments and Consistency messages were most persuasive, though of course their study only considered overall perceived persuasiveness without using a validated scale.

4.2.2. Impact of Message Types on Quality

According to **Figure 3**, for healthy eating messages ARGUMENT FROM EXPERT OPINION WITH GOAL was the highest rated in quality while ARGUMENT FROM POSITION TO KNOW WITH GOAL was the lowest. There was a significant effect of message type on Quality [$F_{(4, 244)} = 12.14, p < 0.001$]. There was a significant difference ($p < 0.05$) between:

1. ARGUMENT FROM POSITION TO KNOW WITH GOAL and the other message types,
2. ARGUMENT FROM SUNK COST WITH ACTION and the other message types except PRACTICAL REASONING WITH GOAL,
3. PRACTICAL REASONING WITH GOAL and the other message types except ARGUMENT FROM COMMITMENT WITH GOAL, and
4. ARGUMENT FROM COMMITMENT WITH GOAL and the other message types except ARGUMENT FROM EXPERT OPINION WITH GOAL.

Table 7 shows the homogeneous subsets. This partially supports the hypothesis (H1) that perceived persuasiveness on each factor differs for different message types. However, it should be noted that one Authority message is the worst and one the best on Quality. This may either be caused by attributes of the message itself, or by one of the Authority argumentation schemes resulting in higher quality messages than the other one.

4.2.3. Impact of Message Types on Capability

According to **Figure 3**, ARGUMENT FROM EXPERT OPINION WITH GOAL was slightly higher rated in quality compared to the other message types. There was no significant effect of message type on Capability [$F_{(4, 244)} = 0.98, p > 0.05$]. **Table 7** shows the homogeneous subsets. This does not support the hypothesis (H1) that perceived persuasiveness of each factor differs for different

message types. All message types performed equally well on Capability, which was above the midpoint of the scale.

4.2.4. Impact of Message Types on Overall Perceived Persuasiveness

According to **Figure 3**, ARGUMENT FROM COMMITMENT WITH GOAL was the highest rated overall while ARGUMENT FROM POSITION TO KNOW WITH GOAL was the lowest. There was a significant effect of message type on Overall Perceived Persuasiveness [$F_{(4, 244)} = 4.98, p < 0.01$]. ARGUMENT FROM POSITION TO KNOW WITH GOAL was significantly different from ARGUMENT FROM EXPERT OPINION WITH GOAL and ARGUMENT FROM COMMITMENT WITH GOAL ($p < 0.05$). The rest were non-significant. **Table 8** shows the homogeneous subsets. This partially supports the hypothesis (H1) that each factor differs on different message types.

4.3. Study 2: Validation of the Perceived Persuasiveness Scale

To determine the construct validity of the developed scale in Study 1 and replicate the scale-testing, we:

1. Used an 80-20 split validation on the original dataset of Study 1. With this specific combination, the developed scale resulted in an acceptable model fit for 80% (TLI = 0.975, CFI = 0.985, RMSEA = 0.081) and 20% of the data (TLI = 0.975, CFI = 0.985, RMSEA = 0.080).
2. Used the dataset obtained from the validation in Study 2. With this dataset, the developed model resulted in an acceptable fit (TLI = 0.984, CFI = 0.990, RMSEA = 0.071).

This answers research question RQ2, validating the scale.

4.4. Study 2: Impact of Message Types on Factors

Figure 4 shows the mean Effectiveness, Quality, Capability, and Overall perceived persuasiveness of message types used for email security messages. As before, the Overall perceived persuasiveness was calculated as the mean of the factors Effectiveness, Quality, and Capability.

A one-way repeated measures MANOVA with Effectiveness, Quality, Capability, and Overall perceived persuasiveness as dependent variables and message type as the independent variable provided the results for the analyses given below. To determine the homogeneous subsets, the Ryan-Einot-Gabriel-Welsch Range was selected as *post-hoc* test since we have more than 3 levels within the independent variable (i.e., message type).

4.4.1. Impact of Message Types on Effectiveness

According to **Figure 4**, ARGUMENT FROM EXPERT OPINION WITH GOAL was the highest rated in Effectiveness while ARGUMENT FROM COMMITMENT WITH GOAL was the lowest. There was a significant effect of message type on Effectiveness [$F_{(4, 568)} = 4.77, p < 0.01$]. ARGUMENT FROM COMMITMENT WITH GOAL was significantly different from ARGUMENT FROM POSITION TO KNOW WITH GOAL and ARGUMENT FROM EXPERT OPINION WITH GOAL ($p < 0.05$). The rest were non-significant. **Table 9** shows the homogeneous subsets. This partly

supports hypothesis H2, namely that perceived persuasiveness in terms of Effectiveness differs for different message types.

The subsets show that Authority messages in the email security domain performed better on Effectiveness than Commitments and Consistency messages. This is in line with the findings of the study by Thomas et al. (2017) and contradicts what was found in Study 1 for the healthy eating messages.

4.4.2. Impact of Message Types on Quality

According to **Figure 4**, ARGUMENT FROM EXPERT OPINION WITH GOAL was the highest rated in Quality while ARGUMENT FROM COMMITMENT WITH GOAL was the lowest. There was a significant effect of message type on Quality [$F_{(4, 568)} = 11.97, p < 0.001$]. ARGUMENT FROM EXPERT OPINION WITH GOAL was significantly different from the other message types ($p < 0.05$). The rest were non-significant. **Table 9** shows the homogeneous subsets. This partially supports hypothesis H3, namely that perceived persuasiveness in terms of Quality differs for different message types.

We observe that ARGUMENT FROM EXPERT OPINION WITH GOAL was rated significantly higher than the other message

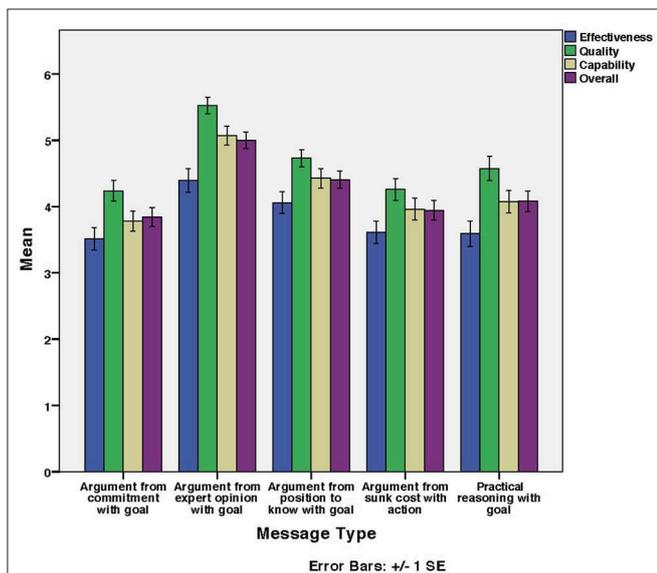


FIGURE 4 | Email security messages: Mean of factors' and overall ratings for developed scale per message type.

types and that the other Authority message had the second highest mean. Therefore, in the domain of email security, we can conclude that principle of Authority seems most persuasive when considering Quality. We note that ARGUMENT FROM EXPERT OPINION WITH GOAL performed best on Quality in both Studies, so this argumentation scheme seems to result in good quality messages. In contrast, ARGUMENT FROM POSITION TO KNOW WITH GOAL did not do as well in the healthy eating domain. It is possible that this is a domain effect, with people trusting people with experience more in the cyber-security domain than in the healthy eating domain. We will investigate this finding further as future work.

4.4.3. Impact of Message Types on Capability

According to **Figure 4**, ARGUMENT FROM EXPERT OPINION WITH GOAL was the highest rated in Capability while ARGUMENT FROM COMMITMENT WITH GOAL was the lowest. There was a significant effect of message type on Capability [$F_{(4, 568)} = 10.84, p < 0.001$]. There was significant difference ($p < 0.05$) between

1. ARGUMENT FROM EXPERT OPINION WITH GOAL and the other message types.
2. ARGUMENT FROM COMMITMENT WITH GOAL and ARGUMENT FROM POSITION TO KNOW WITH GOAL.

There were no significant differences between ARGUMENT FROM SUNK COST WITH ACTION and PRACTICAL REASONING WITH GOAL. **Table 9** shows the homogeneous subsets. This partially supports hypothesis H4 that perceived persuasiveness in terms of Capability differs for different message types.

We observe that ARGUMENT FROM EXPERT OPINION WITH GOAL was rated significantly higher than other message types, and that the other Authority message was rated second highest. Therefore, we can conclude that the principle of Authority was also most persuasive when considering Capability. Again, we can see domain effects in this finding, with ARGUMENT FROM POSITION TO KNOW performing better compared to other message types in the email security domain.

4.4.4. Impact of Message Types on Overall Perceived Persuasiveness

According to **Figure 4**, ARGUMENT FROM EXPERT OPINION WITH GOAL was the highest rated in overall perceived persuasiveness whilst ARGUMENT FROM COMMITMENT WITH

TABLE 9 | Study 2: Homogeneous subsets for Effectiveness, Quality, and Capability.

Message type	Effectiveness			Quality			Capability			
	Mean			Mean			Mean			
	N	S1	S2	N	S1	S2	N	S1	S2	S3
Argument from commitment with goal	115	3.51		115	4.24		115	3.78		
Practical reasoning with goal	111	3.59		115	4.26		115	3.96	3.96	
Argument from sunk cost with action	115	3.61		111	4.58		111	4.08	4.08	
Argument from position to know with goal	117	4.06	4.06	117	4.73		117		4.43	
Argument from expert opinion with goal	115		4.40	115		5.52	115			5.07

TABLE 10 | Study 2: Homogeneous subsets for overall perceived persuasiveness.

Message type	N	Mean		
		Subset 1	Subset 2	Subset 3
Argument from commitment with goal	115	3.84		
Argument from sunk cost with action	115	3.94	3.94	
Practical reasoning with goal	111	4.08	4.08	
Argument from position to know with goal	117		4.41	
Argument from expert opinion with goal	115			5.00

GOAL was the lowest. There was a significant effect of message type on overall perceived persuasiveness [$F_{(4, 568)} = 11.24, p < 0.001$]. **Table 10** shows the homogeneous subsets. This partially supports hypothesis H5 that the overall perceived persuasiveness differs for different message types.

The overall perceived persuasiveness results are similar to those for “Impact of message type on Capability”; again overall Authority messages performed well, and better than in the healthy eating domain.

5. DISCUSSION

Our studies resulted in a validated perceived persuasiveness scale as well as insights into the perceived persuasiveness of different message types.

5.1. The Perceived Persuasiveness Scale

Regarding the scale, as mentioned in the limitations of the systematic literature review, there are some other papers that proposed persuasiveness scales that were not part of the review. The uptake of these scales has been limited as judged by them not having been used in the reviewed papers. However, it is interesting to see how these scales compare to the one developed in this paper, and to consider what overlap/differences there are.

First, Feltham (1994) developed and validated a Persuasive Discourse Inventory (PDI) scale based on Aristotle’s three types of persuasion: ethos, pathos, and logos (see **Table 11**). Ethos relates to the credibility of the message source, pathos to the message’s affective appeal, and logos to its rational appeal. To validate the PDI scale, they mainly considered Cronbach’s alpha rather than conducting a factor analysis as was done in this paper. Their results suggest that there may be cross-loadings between their scale factors as they found a positive correlation between Logos and Ethos. They also did not consider whether the scale performed well across domains, as their reassessment was conducted in a very similar domain. Regarding the scale content, the scale developed in this paper has more items that directly inquire into a message’s perceived persuasiveness rather than the emotional and logical elements present in the messages, though Ethos, Logos, and Pathos still play a role. Several Ethos related items were included in our initial scale development items, namely trustworthy, believable, and credible. One of these items (cf. trustworthy) has remained in the validated scale as part of the Quality factor. The “accurate” item that is part of

TABLE 11 | Persuasive Discourse Inventory (Feltham, 1994).

Ethos scale items: Ethos = E1+E2+E3+E4+E5 (range: 5-35)

- E1) unbelievable / believable
- E2) not credible / credible
- E3) not trustworthy / trustworthy
- E4) unreliable / reliable
- E5) undependable / dependable

Logos scale items: Logos = L1+L2+L3+L4+L5 (range 5-35)

- L1) not rational / rational
- L2) not informative / informative
- L3) does not deal with facts / deals with facts
- L4) not knowledgeable / knowledgeable
- L5) not logical / logical

Pathos scale items: Pathos = P1+P2+P3+P4+P5+P6+P7 (range: 7-49)

- P1) does not affect my feelings / affects my feelings
- P2) does not touch me emotionally / touches me emotionally
- P3) is not stimulating / is stimulating
- P4) does not reach out to me / reaches out to me
- P5) is not stirring / is stirring
- P6) is not moving / is moving
- P7) is not exciting / is exciting

the Quality factor can be interpreted as on the overlap between Ethos and Logos, as it on the one hand gives a sense of being reliable, and on the other of being based on facts/rational/logical. Regarding Pathos, the item “This message has the potential to inspire users” in the Capability factor is clearly related to Pathos (as was the item “motivating” that did not make it into the final scale).

Second, Lehto et al. (2012) developed a model with factors that predict perceived persuasiveness, and as part of this also considered the internal consistency of items to measure these factors. Several of their factors (e.g., dialogue support, design aesthetics) are not directly about persuasive messages *per se* but rather about the overarching behavioral intervention system they were studying. The aim of their work was not to develop a scale, so they did not try to develop factors that are independent of each other, but were mainly interested in how the factors related to each other. In fact, despite finding adequate internal consistency, they found quite a lot of cross-loadings, with items from one factor loading above 0.5 on other factors as well. Their validation was only in the health domain, and many of their questions specifically related to their intervention (e.g., a primary task support item “NIV provides me with a means to lose weight,” a dialogue support item “NIV provides me with appropriate counseling,” a perceived credibility item “NIV is made by health professionals”). So, this work did not result in a multi independent factors scale that can be used in multiple domains, like the scale developed in this paper. Considering the factors they considered, Perceived Credibility overlaps with the Quality factor in our scale (cf. trustworthy). Primary Task

support is related to the Effectiveness factor in our scale (e.g., “helps me change [my behavior]” is related to “causes a change in my behavior”). Their Perceived Persuasiveness factor has some relation to our Capability factor (e.g., compare “has an influence on me” and “has the potential to influence user behavior,” “makes me reconsider [my behavior],” and “has the potential to change user behavior”).

Third, Allen et al. (2000) compared the persuasiveness of statistical and narrative evidence in a message, and produced two scales to perform this study: a Credibility scale (measuring the extent to which one trusts the message writer) and an Attitude scale (measuring the extent to which one accepts the message’s conclusion). They checked that each scale only contained one factor, and that each scale was internally consistent (in terms of Cronbach’s alpha). They did not, however, consider whether items from one scale cross-loaded onto the other scale (e.g. the items “I think the writer is wrong” from the Attitude scale and “the writer is dishonest” from the Credibility scale seem related, so cross-loadings may well occur). They also did not remove an item with low factor loading (“the writing style is dynamic,” loading 0.40) from the Credibility scale, which may indicate a poor scale structure (MacCallum et al., 1999). Their scales only measure some aspects of persuasiveness; for example, they do not measure the message’s potential to inspire, or to cause behavior change.

Fourth, Popova et al. (2014), Jasek et al. (2015), and Yzer et al. (2015) used multi-item scales, but without a development phase. Popova et al. (2014) used five items (convincing-unconvincing, effective-ineffective, believable-unbelievable, realistic-unrealistic, and memorable-not memorable), Jasek et al. (2015) 13 (boring, confusing, convincing, difficult to watch, informative, made me want to quit smoking, made me want to smoke, made me stop and think, meaningful to me, memorable, powerful, ridiculous, terrible), and Yzer et al. (2015) 7 (convincing, believable, memorable, good, pleasant, positive, for someone like me). There is considerable overlap between these items and the ones we used for the scale development, though there are some items in these papers that seem more related to usability (e.g., “confusing”) and some more related to feelings (e.g., “pleasant,” “terrible”).

Fifth, McLean et al. (2016) developed a scale from 13 items for measuring the persuasiveness of messages to reduce stigma about bulimia. They only performed an exploratory factor analysis (using ratings of only 10 messages), so no real validation. Their scale has two factors; one they describe as convincingness and the other as likelihood of changing attitudes toward bulimia. The first factor includes items such as “believable” and “convincing,” which were part of our initial items for scale development and are related to the Quality factor in our scale. The second factor is related to the Capability factor of our scale.

In summary, the scale developed in this paper is unique in that it was developed from a large set of items covering a wide range of aspects of persuasiveness, was developed and validated across two domains, and has been shown to consist of three independent factors, with good internal consistency. The comparison of scale content with the content of other scales shows that the scale also provides reasonable coverage of concepts deemed important in

the literature (for example, some aspects of Ethos, Pathos, and Logos are present).

5.2. Persuasiveness of Message Types

As a side effect of our studies, we also gained insights into the persuasiveness of message types. There have been several other papers investigating this, though these studies have only investigated the impact of Cialdini’s principles and not the finer-grained argumentation schemes. For instance, Orji et al. (2015) and Thomas et al. (2017) investigated the persuasiveness of Cialdini’s principles for healthy eating, Smith et al. (2016) for reminders to cancer patients, Ciocarlan et al. (2018) for encouraging small acts of kindness, and Oyibo et al. (2017) in general without mentioning specific domains.

Thomas et al. (2017) found that Authority messages were most persuasive and Liking least persuasive. Orji et al. (2015) found that Commitment and Reciprocity were the most persuasive over all ages and gender, whereas Consensus and Scarcity were the least persuasive. They found that females responded better to Reciprocity, Commitment, and Consensus messages than males. They also observed that adults responded better to Commitment than younger adults, and younger adults responded better to Scarcity than adults. Smith et al. (2016) observed that Authority and Liking were the most popular for the first reminder, and there was a preference for using Scarcity and Commitment for the second reminder. Ciocarlan et al. (2018) found that the Scarcity message worked best. Oyibo et al. (2017) observed that their participants were more susceptible to Authority, Consensus, and Liking.

The conflicting results of these studies can have several causes. Firstly, the studies were conducted in different domains. Our studies in this paper have shown that the persuasiveness of message types is in fact domain dependent. For example, we found in the Healthy Eating domain that some of the Authority-linked argumentation schemes scored badly on Effectiveness, and one of them was also worst on persuasiveness overall, whilst in the Email Security domain Authority-linked argumentation schemes scored best. Secondly, the studies used very different (and not validated) ways of measuring persuasiveness. So, it would be interesting to repeat all of these studies in a variety of domains using the scale developed in this paper. Thirdly, these studies did not consider the finer-grained argumentation schemes, but only Cialdini’s principles. It is possible that, for example, the Authority messages used in one study followed a different argumentation scheme (within the Authority set) than those in another study. Finally, in contrast to our studies, none of these papers considered the individual factors of persuasiveness, but only considered persuasiveness as a whole. Our studies show that it is possible for a message type to score badly on one dimension on persuasiveness whilst scoring well on the others.

In summary, the most important results in this paper regarding the persuasiveness of message types are that (1) this persuasiveness is domain dependent, (2) investigating the finer-grained argumentation schemes matters as different results can be obtained for different argumentation schemes that are linked to the same Cialdini’s principles, and (3) investigating the

different factor of persuasiveness matters as different results can be obtained for the different factors.

6. CONCLUSIONS

In this paper, we developed and validated a perceived persuasiveness scale to be used when conducting studies on digital behavior interventions. We conducted two studies in different domains to develop and validate this scale, namely in the healthy eating domain and the email security domain. The validated scale has 3 factors (Effectiveness, Quality, and Capability) and 9 scale items as illustrated in **Table 6**. We also discussed how this scale relates to and extends on earlier work on persuasiveness scales.

In addition to developing a scale, and to show its usefulness, we analyzed the impact of message types on the different developed scale factors. We found that message type significantly impacts on Effectiveness, Quality, and overall perceived persuasiveness in studies in both the healthy eating and email security domains. We also found a significant impact of message type on Capability in the email security domain. The three factors (as shown in the validation) measure different aspects of perceived persuasiveness. One example where this can also be seen is for the ARGUMENT FROM EXPERT OPINION WITH GOAL message type, which performs relatively badly on Effectiveness in the healthy eating domain but well on Quality in that domain. The persuasiveness of messages is clearly domain dependent. Additionally, our studies show that it is worthwhile to investigate the finer-grained argumentation schemes rather than just Cialdini's principles. We discussed related work on measuring the persuasiveness of message types and explained the conflicting findings in those studies.

As shown in our literature review, researchers working on digital behavior interventions tend to use their own scales, without proper validation of those scales, to investigate perceived persuasiveness. The validated scale developed in this paper can be used to improve such studies and will make it easier to compare the results of different studies and in different domains. We plan

REFERENCES

- Allen, M., Bruflat, R., Fucilla, R., Kramer, M., McKellips, S., Ryan, D. J., et al. (2000). Testing the persuasiveness of evidence: combining narrative and statistical forms. *Commun. Res. Rep.* 17, 331–336. doi: 10.1080/08824090009388781
- Anagnostopoulou, E., Magoutas, B., Bothos, E., Schrammel, J., Orji, R., and Mentzas, G. (2017). "Exploring the links between persuasion, personality and mobility types in personalized mobility applications," in *Persuasive Technology: Development and Implementation of Personalized Technologies to Change Attitudes and Behaviors*, eds P. W. de Vries, H. Oinas-Kukkonen, L. Siemons, N. Beerlage-de Jong, and L. van Gemert-Pijnen (Cham: Springer International Publishing), 107–118.
- Busch, M., Patil, S., Regal, G., Hochleitner, C., and Tscheligi, M. (2016). "Persuasive information security: techniques to help employees protect organizational information security," in *Persuasive Technology*, eds A. Meschtscherjakov, B. De Ruyter, V. Fuchsberger, M. Murer, and M. Tscheligi (Cham: Springer International Publishing), 339–351.
- Busch, M., Schrammel, J., and Tscheligi, M. (2013). *Personalized Persuasive Technology – Development and Validation of Scales for Measuring Persuadability*. Berlin; Heidelberg: Springer, 33–38.
- Chang, J.-H., Zhu, Y.-Q., Wang, S.-H., and Li, Y.-J. (2018). Would you change your mind? an empirical study of social impact theory on facebook. *Telem. Inform.* 35, 282–292. doi: 10.1016/j.tele.2017.11.009
- Cialdini, R. B. (2009). *Influence: The Psychology of Persuasion*. New York, NY: HarperCollins e-books.
- Ciocarlan, A., Masthoff, J., and Oren, N. (2018). "Kindness is contagious: study into exploring engagement and adapting persuasive games for wellbeing," in *Proceedings of the 26th Conference on User Modeling, Adaptation and Personalization, UMAP '18* (New York, NY: ACM), 311–319.
- Cook, A., Pryer, J., and Shetty, P. (2000). The problem of accuracy in dietary surveys. Analysis of the over 65 UK national diet and nutrition survey. *J. Epidemiol. Commun. Health* 54, 611–616. doi: 10.1136/jech.54.8.611
- Dziuban, C. D., and Shirkey, E. C. (1980). Sampling adequacy and the semantic differential. *Psychol. Rep.* 47, 351–357. doi: 10.2466/pr0.1980.47.2.351
- Feltham, T. S. (1994). Assessing viewer judgement of advertisements and vehicles: scale development and validation. *ACR North Am. Adv.* 21, 531–535.
- Grasso, F., Cawsey, A., and Jones, R. (2000). Dialectical argumentation to solve conflicts in advice giving: a case study in the promotion of healthy nutrition. *Int. J. Hum. Comput. Stud.* 53, 1077–1115. doi: 10.1006/ijhc.2000.0429

to use the scale to study the impact of message personalization across domains.

The work presented in this paper has several limitations. Firstly, we validated the scale in two domains (healthy eating and email security), and this validation needs to be extended to more domains. Secondly, the scale reliability needs to be verified. To investigate this, we need to perform a test-retest experiment in which participants complete the same scale on the same items twice, with an interval of several days between the two measurements. This also would need to be done in multiple domains. Thirdly, we need to repeat our studies into the impact message types with more messages and in more domains.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by CoPs ethics committee University of Aberdeen. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

RT and JM contributed to the conception and design of the study. RT implemented the study, performed the statistical analysis, and wrote the first draft of the manuscript. All authors wrote sections of the manuscript, contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This work on cyber-security in this manuscript was supported by the EPSRC under Grant EP/P011829/1.

- Ham, C.-D., Nelson, M. R., and Das, S. (2015). How to measure persuasion knowledge. *Int. J. Advertis.* 34, 17–53. doi: 10.1080/02650487.2014.994730
- Hammer, S., Lugin, B., Bogomolov, S., Janowski, K., and André, E. (2016). “Investigating politeness strategies and their persuasiveness for a robotic elderly assistant,” in *Persuasive Technology*, eds A. Meschtscherjakov, B. De Ruyter, V. Fuchsberger, M. Murer, and M. Tscheligi (Cham: Springer International Publishing), 315–326.
- Hossain, M. T., and Saini, R. (2014). Suckers in the morning, skeptics in the evening: time-of-day effects on consumers’ vigilance against manipulation. *Market. Lett.* 25, 109–121. doi: 10.1007/s11002-013-9247-0
- Howitt, D., and Cramer, D. (2014). *Introduction to SPSS Statistics in Psychology*. Pearson Education.
- Hu, L., and Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equat. Model. A Multidiscipl. J.* 6, 1–55. doi: 10.1080/10705519909540118
- Jasek, J. P., Johns, M., Mbamalu, I., Auer, K., Kilgore, E. A., and Kansagra, S. M. (2015). One cigarette is one too many: evaluating a light smoker-targeted media campaign. *Tobacco Control* 24, 362–368. doi: 10.1136/tobaccocontrol-2013-051348
- Kaptein, M., Markopoulos, P., de Ruyter, B., and Aarts, E. (2009). *Can You Be Persuaded? Individual Differences in Susceptibility to Persuasion*. Berlin; Heidelberg: Springer, 115–118.
- Koch, T., and Zerbac, T. (2013). Helpful or harmful? How frequent repetition affects perceived statement credibility. *J. Commun.* 63, 993–1010. doi: 10.1111/jcom.12063
- Lehto, T., Oinas-Kukkonen, H., and Drozd, F. (2012). “Factors affecting perceived persuasiveness of a behavior change support system,” in *Thirty Third International Conference on Information Systems, Orlando*. Orlando.
- MacCallum, R. C., Widaman, K. F., Zhang, S., and Hong, S. (1999). Sample size in factor analysis. *Psychol. Methods* 4:84.
- MacKenzie, S. B., and Lutz, R. J. (1989). An empirical examination of the structural antecedents of attitude toward the ad in an advertising pretesting context. *J. Market.* 53, 48–65.
- Mazzotta, I., de Rosis, F., and Carofiglio, V. (2007). Portia: a user-adapted persuasion system in the healthy-eating domain. *IEEE Intell. Syst.* 22, 42–51. doi: 10.1109/MIS.2007.115
- McLean, S. A., Paxton, S. J., Massey, R., Hay, P. J., Mond, J. M., and Rodgers, B. (2016). Identifying persuasive public health messages to change community knowledge and attitudes about bulimia nervosa. *J. Health Commun.* 21, 178–187. doi: 10.1080/10810730.2015.1049309
- Meschtscherjakov, A., Gärtner, M., Mirnig, A., Rödel, C., and Tscheligi, M. (2016). “The persuasive potential questionnaire (PPQ): Challenges, drawbacks, and lessons learned,” in *Persuasive Technology*, eds A. Meschtscherjakov, B. De Ruyter, V. Fuchsberger, M. Murer, and M. Tscheligi (Cham: Springer International Publishing), 162–175.
- Oduor, M., and Oinas-Kukkonen, H. (2017). “Commitment devices as behavior change support systems: a study of users’ perceived competence and continuance intention,” in *Persuasive Technology: Development and Implementation of Personalized Technologies to Change Attitudes and Behaviors*, eds P. W. de Vries, H. Oinas-Kukkonen, L. Siemons, N. Beerlage-de Jong, and L. van Gemert-Pijnen (Cham: Springer International Publishing), 201–213.
- O’Keefe, D. J. (2018). Message pretesting using assessments of expected or perceived persuasiveness: evidence about diagnosticity of relative actual persuasiveness. *J. Commun.* 68, 120–142. doi: 10.1093/joc/jqx009
- Orji, R. (2014). “Exploring the persuasiveness of behavior change support strategies and possible gender differences,” in *Conference of 2nd International Workshop on Behavior Change Support Systems*, Vol. 1153, eds L. van Gemert-Pijnen, S. Kelders, A. Oorni, and H. Oinas-Kukkonen (Aachen: CEUR-WS), 41–57.
- Orji, R., Mandryk, R. L., and Vassileva, J. (2015). “Gender, age, and responsiveness to cialdini’s persuasion strategies,” in *Persuasive Technology*, eds T. MacTavish and S. Basapur (Cham: Springer International Publishing), 147–159.
- Orji, R., Vassileva, J., and Mandryk, R. L. (2014). Modeling the efficacy of persuasive strategies for different gamer types in serious games for health. *User Model. User Adapt. Interact.* 24, 453–498. doi: 10.1007/s11257-014-9149-8
- Oyibo, K., Orji, R., and Vassileva, J. (2017). “Investigation of the influence of personality traits on cialdini’s persuasive strategies,” in *Proceedings of the 2nd International Workshop on Personalization in Persuasive Technology*, Vol. 1833, eds R. Orji, M. Reisinger, M. Busch, A. Dijkstra, M. Kaptein, and E. Mattheiss (CEUR-WS), 8–20.
- Popova, L., Neilands, T. B., and Ling, P. M. (2014). Testing messages to reduce smokers’ openness to using novel smokeless tobacco products. *Tobacco Control* 23, 313–321. doi: 10.1136/tobaccocontrol-2012-050723
- Purpura, S., Schw, V., Williams, K., Stubler, W., and Sengers, P. (2011). “Fit4life: The design of a persuasive technology promoting healthy behavior and ideal weight,” in *Proceedings of the International Conference on Human Factors in Computing Systems, CHI 2011, Vancouver, BC, Canada, May 7-12, 2011* (Vancouver: ACM), 423–432.
- Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., and King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: a review. *J. Educ. Res.* 99, 323–338. doi: 10.3200/JOER.99.6.323-338
- Smith, K. A., Dennis, M., and Masthoff, J. (2016). “Personalizing reminders to personality for melanoma self-checking,” in *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization* (Halifax), 85–93.
- Thomas, R. J., Masthoff, J., and Oren, N. (2017). “Adapting healthy eating messages to personality,” in *Persuasive Technology. 12th International Conference, PERSUASIVE 2017, Proceedings* (Amsterdam: Springer), 119–132.
- Thomas, R. J., Oren, N., and Masthoff, J. (2018). “ArguMessage: a system for automation of message generation using argumentation schemes,” in *Proceedings of AISB Annual Convention 2018, 18th Workshop on Computational Models of Natural Argument* (Liverpool), 27–31.
- Walton, D., Reed, C., and Macagno, F. (2008). *Argumentation Schemes*. New York, NY: Cambridge University Press.
- Wells, S., Kotkanen, H., Schläfli, M., Gabrielli, S., Masthoff, J., Jylhä, A., et al. (2014). Towards an applied gamification model for tracking, managing, & encouraging sustainable travel behaviours. *EAI Endors. Trans. Ambient Syst.* 1:e2. doi: 10.4108/amsys.14.e2
- Yzer, M., LoRusso, S., and Nagler, R. H. (2015). On the conceptual ambiguity surrounding perceived message effectiveness. *Health Commun.* 30, 125–134. doi: 10.1080/10410236.2014.974131
- Zhang, K. Z., Zhao, S. J., Cheung, C. M., and Lee, M. K. (2014). Examining the influence of online reviews on consumers’ decision-making: a heuristic systematic model. *Decis. Support Syst.* 67, 78–89. doi: 10.1016/j.dss.2014.08.005
- Zhao, X., Strasser, A., Cappella, J. N., Lerman, C., and Fishbein, M. (2011). A measure of perceived argument strength: a reliability and validity. *Commun. Methods Meas.* 5, 48–75. doi: 10.1080/19312458.2010.547822

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.

The reviewer KS declared a past collaboration with one of the authors JM to the handling editor.

Copyright © 2019 Thomas, Masthoff and Oren. 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.