



# Krauzlis' Strange Inversion of Reasoning

**Domenico Guarino\***

Centre National de la Recherche Scientifique, Unité de Neurosciences Information et Complexité, Paris, France

**Keywords:** attention, multiple-draft model, homunculus, cell assembly, basal ganglia, superior colliculus

“Attention” is a central faculty of the mind (Baars, 1988; Minsky, 1988). It has been early proposed that it is a process selecting some of the available information to focus on enhanced processing (Deutsch and Deutsch, 1963; Treisman and Gelade, 1980). Contrasting this approach, Richard J. Krauzlis recently (2014; 2017) proposed that attention is an *effect* of selection processes instead of being their *cause*. In this article, I will argue that Krauzlis' view realizes what Dennett (2009) calls a “strange inversion of reasoning,” in which the competence of selecting behaviors does not need the comprehension of any attentive homunculus. In order to make the equivalence, I will detail an interface mapping between a philosophical position (Dennett's Multiple-Draft model, 1991) and a neuroscientific position (Krauzlis' attention as an effect). Such a mapping is meant to encourage a scientifically grounded dialogue between the two disciplines.

According to the original feature-integration theory of visual attention, “features are registered early, automatically, and in parallel across the visual field, while objects are identified separately and only at a later stage, which requires focused attention” (Treisman and Gelade, 1980, p. 98).

The selection mechanism of attention, and its corresponding neural correlates, have been henceforth the target of further research. Crick (1984) proposed, with his “searchlight” hypothesis, that the thalamic reticular nucleus controls the focus of attention, whose content is represented at the cortical level. In fact, reticular cells receive connections from both feedforward thalamic and feedback cortical pathways, and they also inhibit each other, thus having the necessary circuitry to implement competition among their afferents (Sherman and Guillery, 2001). Over time, several other brain areas have been implicated in attentive behavior (Krauzlis et al., 2013): sensory and parietal cortex, frontal cortices like the frontal eye field, with a growing role for the superior colliculus (SC). Recently, Lovejoy and Krauzlis (2010) locally and reversibly suppressed the SC of monkeys performing a visual detection task. For stimuli placed in the affected part of SC, monkeys' performance was drastically reduced. Using the searchlight hypothesis, the conclusion would be that SC selects which part of the visual field has to be the target of attention, by supplying the ascending pathway to cortex with a saliency map (Koch and Ullman, 1987). Thus, the next step (Zénon and Krauzlis, 2012) consisted in transiently inactivating the SC during a detection task and measuring responses in corresponding visual cortical areas, with the prediction that cortical activity would be reduced due to SC inactivation. Surprisingly, the inactivation of SC did impair task performance but did not reduce cortical activity, challenging the idea that SC selects the target of visual attention.

Instead of searching for another correlate center of attention, Krauzlis and colleagues proposed a different model: *attention is an effect rather than a cause*. It would be the functional consequence of competition between “interpretations of the current state by the basal ganglia rather than competition to determine how sensory data is represented in the neocortex” (Krauzlis et al., 2014, p. 458). Such a reversal of how attention is intended realizes a “strange inversion of reasoning” (Dennett, 2009), where processes that do not have comprehension can nonetheless show competence. In Krauzlis case, attentive behaviors would not result from the comprehension of a homunculus selecting which sensory data is represented in neocortex. They would instead result from the competence of selective processes in the basal ganglia.

## OPEN ACCESS

### Edited by:

James W. Grau,  
Texas A&M University, United States

### Reviewed by:

Colin Allen,  
University of Pittsburgh, United States  
Adam R. Ferguson,  
University of California, San Francisco,  
United States

### \*Correspondence:

Domenico Guarino  
domenico.guarino@gmail.com

**Received:** 25 April 2018

**Accepted:** 29 June 2018

**Published:** 20 July 2018

### Citation:

Guarino D (2018) Krauzlis' Strange  
Inversion of Reasoning.  
Front. Syst. Neurosci. 12:34.  
doi: 10.3389/fnsys.2018.00034

Further exploring this link prompts the identification of an interface—revolving around the concept of cell assembly (Hebb, 1949)—yielding a synergistic mapping between Dennett’s Multiple-Drafts model (MDM) and Krauzlis’ model regarding attention.

Dennett’s MDM fundamental thesis is: “There is no single stream of consciousness, no place where ‘it all comes together’ for a Central Meaner. There are multiple drafts, separate channels in which specialist circuits try to do their various things. Most of these fragmentary drafts of narrative play short-lived roles in the modulation of activity but some get promoted to further functional roles, in swift succession. The seriality of this process is not a hard-wired design feature, but rather the upshot of a succession of coalitions of these specialists” (Dennett, 1991, p. 253). These four sentences are the core building blocks of Dennett’s explanation of consciousness. By analyzing each sentence, I will now build the interface and use it to map the two models (see **Table 1** for a direct comparison).

The first sentence rejects the homunculus, and Krauzlis shares it: “Viewing attention as an effect rather than a cause has the advantage of eliminating the need to find the homunculus that aims the spotlight of attention. Instead, we should be able to explain the functional circuits that give rise to the phenomenology of attention without using the word ‘attention’” (Krauzlis et al., 2014, p. 461).

In the second sentence, Dennett was intentionally vague using the term “draft.” A possible denotation for draft, however, may be the Hebbian cell assembly: “a diffuse structure comprising cells in the cortex and diencephalon (and also, perhaps, in the basal ganglia of the cerebrum), capable of acting briefly as a closed system, delivering facilitation to other such systems and usually having a specific motor facilitation” (Hebb, 1949, p. xix). As explicitly put in Hebb’s definition, the concept of cell assembly identifies sets of cells, spanning the whole sensation-action arc. However, over time, two overlapping versions of the same concept emerged. *Object-action cell assemblies*, representing potential associations between sensory objects and actions (Humphreys and Riddoch, 2000), and (generic) *cell assemblies*, with a broader sense of “set[s] of neurons among which excitatory connections have been potentiated” (Harris, 2005). Therefore, if we denote Dennett’s “specialist circuit” by a set of functionally related areas (e.g., visual: lateral geniculate nucleus, striate and extrastriate cortices, SC, ..., as in Akins, 1996), then a “draft” can be denoted by an active (generic) cell assembly—one out of many that could be active at any given moment in the “specialist.”

The third sentence fits the proposed interface well, highlighting the fluidity of cell assemblies, and introducing a selection process involving them: “some get promoted to further functional roles, in swift succession.” But how, exactly?

This question is crucial for the MDM, and the interface may offer an answer, provided it succeeds in the terminology mapping. Krauzlis considers that “Good decision making depends crucially on properly identifying the current state of the animal and its environment” (Krauzlis et al., 2014, p. 457). This wording implies a “state” that has to be recognized. The concept of “state” has its root at the beginning of cybernetics, where every physical system (including an organism, its nervous system, and its environment) can be described by a finite set of states and state transitions (Ashby, 1961, p. 10). However, this powerful abstraction has been subsequently challenged because many relations between an organism and its environment, or “affordances” Gibson (2013), are equally possible in the same environmental state. The neural components of such relations have been mapped onto object-action cell assemblies (Humphreys and Riddoch, 2000; Cisek, 2007). Therefore, a distinction has to be made between Krauzlis’ “candidate state”—denoted by any object-action cell assembly that can be potentially enacted—and his “current state” that should be denoted by an object-action cell assembly which has been selected, among many, to be actually enacted.

According to Krauzlis, the “current state” is evaluated by the basal ganglia. In fact, they receive sensory patterns (Voorn et al., 2004) and action plans (Jin et al., 2014) from cortex, threat detections from the amygdala (LeDoux, 2014), memories of action sequences from hippocampus (Jai and Frank, 2015) with their possible outcomes from prefrontal cortex (Donoso et al., 2014), and stimulus saliency from thalamic intralaminar nuclei (Ding et al., 2010; Doig et al., 2014). Each of these areas, in MDM terms, is a specialist producing a temporary draft. Within the basal ganglia, the striatum has the circuitry to implement competition among them (Tepper and Bolam, 2004; Cisek, 2007). However, “Because different sensory inputs and types of knowledge contribute unequally to different states, their influence on perception and action will be limited by the strength of their contribution to the state that currently dominates the competition” (Krauzlis et al., 2014, p. 458). This point is central to build the interface and realize the terminology mapping, thus it requires further unpacking.

The intrinsic recurrent connections between cells of an assembly stabilize their firing through mutual excitation (Harris, 2005) while also suppressing other assemblies through inhibition

**TABLE 1** | Interface providing a mapping between philosophical and neuroscientific terminology.

Dennett terms	Interface	Krauzlis terms
“Specialist”	Set of brain areas	
“Draft”	(specialist’s) cell assembly	
“Affordance” (Gibson)	(object-action) cell assembly	“Candidate state”
“Coalition of specialists”	Dynamic mutual strengthening between (specialist’s) cell assemblies, eventually leading to an (object-action) cell assembly	Strength of “different sensory inputs and types of knowledge contributions to the current state”
“Winning coalition”	Strengthened (object-action) cell assembly currently dominating the selection process	“Current state”
“Succession of coalitions”	Sequence of (object-action) cell assemblies (or phase sequence)	“Linked chain of states”

(Zenke et al., 2015). The occurrence of a specific assembly in one area both raises the probability of occurrence for some assemblies and quenches the probability of occurrence for some other assemblies in the same or other areas at some later time (Ponzi and Wickens, 2009; Truccolo et al., 2010) by means of short-term synaptic plasticity (von der Malsburg, 1994). Now, two or more specialists, each producing temporary drafts / cell assemblies, can interact through direct or striato-thalamo-cortical connections (Hikosaka et al., 2000; Sommer and Wurtz, 2004, 2006). These interactions, dynamically raising or lowering the probability of occurrence of an assembly given the occurrence of another, are the required mechanism for having dynamic coalitions of specialists' drafts, denoted by mutually strengthening assemblies. Eventually, one of these coalitions temporarily wins the competition in the striatum, and its entailed action is performed. The outcome for the organism is signaled, at the nervous system level, as dopaminergic reward (Threlfell et al., 2012) instrumental in selecting the synapses that will be targeted for growth (Li et al., 2003; Howe et al., 2013; Pelosi and Guarino, 2015), which will shape cell assemblies, and which will turn into (possibly better) behaviors (Montague et al., 1996).

With "state that currently dominates the competition" being the object-action cell assembly currently dominating the selection process in the basal ganglia (Dennett's winning coalition), we may now proceed to the fourth sentence, regarding the seriality and succession of coalitions. The proposed interface still matches: "As circumstances change, a candidate state that provides a poor match in one round of competition could emerge as predominant just moments later. This results in a linked chain of states, where the transition from one to the next is precipitated by some event or change in an internal variable, that gives the new state more support than the preceding one and carries along with a new decision policy" (Krauzlis et al., 2014, p. 458). Here, "candidate state" is denoted, as proposed earlier, by no more than an object-action cell assembly competing with all the others in the striatum. These candidate states, with their associated potential actions, represent different decision policies. Hence, a succession of selected states — comprehension-less but competence-full, thanks to a reinforcement learning process—steers the policy making that we label as "attentional" behavior. The reinforcement behind this succession of states, and its link to attentional behavior, is also fitting the Hebbian "phase sequence" (Hebb, 1949, p. 87).

## REFERENCES

- Akins, K. (1996). Lost the plot? reconstructing Dennett's multiple drafts theory of consciousness. *Mind Lang.* 11, 1–43.
- Ashby, W. R. (1961). *An Introduction to Cybernetics*. New York, NY: Chapman & Hall Ltd.
- Baars, B. J. (1988). *A Cognitive Theory of Consciousness*. New York, NY: Cambridge University Press.
- Cisek, P. (2007). Cortical mechanisms of action selection: the affordance competition hypothesis. *Philos. Trans. R. Soc. B Biol. Sci.* 362, 1585–1599. doi: 10.1098/rstb.2007.2054
- Crick, F. (1984). Function of the thalamic reticular complex: the searchlight hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 81, 4586–4590.

To summarize (see **Table 1**), several cell assemblies (Dennett's "short-lived," "coalitions of specialists," "drafts," or Krauzlis' "candidate states") are dynamically strengthened and compete in the striatum. At every moment, one object-action cell assembly is temporarily dominating the selection process in the striatum (Dennett's winning "coalition," or Krauzlis' "current state"), getting enacted. In time, a series of winning assemblies (Dennett's "succession of coalitions of specialists," or Krauzlis' "linked chain of states") gets selected to be pursued with actions, leading (or not) to rewards for the organism, ultimately promoting the formation, dissolution, and change of cell assemblies. Hence it is not a homunculus, rather learned associations that drive the production of "attentional" behaviors."

The interface I proposed, here, and the mapping between the terminology used by Dennett and that of Krauzlis is meant to encourage the exchange between philosophy of mind and neuroscience. In fact, on the one hand, it seems that the philosophical refinement of questions regarding the mind needs more scientific injections (Dennett, 2003). While, on the other hand, neuroscience is in danger of a technological stale-point because it is lacking more encompassing and purposeful views (Frégnac, 2017). If, and only if, a common terminology will be adopted such an exchange will take place and benefit both disciplines.

## AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and approved it for publication.

## FUNDING

This work was supported by the Centre Nationale de la Recherche Scientifique (CNRS), the national excellency initiative ICode (Lidex Saclay), the French National Agency of Research (ANR Horizontal-V1) and the EC Human Brain Project (HBP, grant agreement no. 604102).

## ACKNOWLEDGMENTS

I would like to express all my gratitude to Daniel C. Dennett, Jan Antolik, and Lyle Muller for reading the original draft and providing insightful comments. This work has been definitely improved by the comments from one reviewer.

- Dennett, D. (2003). Who's on first? heterophenomenology explained. *J. Conscious. Stud.* 10, 19–30.
- Dennett, D. (2009). Darwin's "strange inversion of reasoning." *Proc. Natl. Acad. Sci. U.S.A.* 106(Suppl. 1), 10061–10065. doi: 10.1073/pnas.0904433106
- Dennett, D. C. (1991). *Consciousness Explained*. Boston, MA: Little, Brown and Company.
- Deutsch, J. A. and Deutsch, D. (1963). Attention: some theoretical considerations. *Psychol. Rev.* 70:80.
- Ding, J. B., Guzman, J. N., Peterson, J. D., Goldberg, J. A., and Surmeier, D. J. (2010). Thalamic gating of corticostriatal signaling by cholinergic interneurons. *Neuron* 67, 294–307. doi: 10.1016/j.neuron.2010.06.017
- Doig, N. M., Magill, P. J., Apicella, P., Bolam, J. P., and Sharott, A. (2014). Cortical and thalamic excitation mediate the multiphasic responses of striatal

- cholinergic interneurons to motivationally salient stimuli. *J. Neurosci.* 34, 3101–3117. doi: 10.1523/JNEUROSCI.4627-13.2014.
- Donoso, M., Collins, A. G., and Koehlin, E. (2014). Foundations of human reasoning in the prefrontal cortex. *Science* 344, 1481–1486. doi: 10.1126/science.1252254
- Frégnac, Y. (2017). Big data and the industrialization of neuroscience: a safe roadmap for understanding the brain? *Science* 358, 470–477. doi: 10.1126/science.aan8866
- Gibson, J. J. (2013). “The theory of affordances,” in *The Ecological Approach to Visual Perception* (New York, NY: Taylor and Francis), 119–135.
- Harris, K. D. (2005). Neural signatures of cell assembly organization. *Nat. Rev. Neurosci.* 6, 399–407. doi: 10.1038/nrn1669
- Hebb, D. (1949). *The Organization of Behavior*. New York, NY: Wiley.
- Hikosaka, O., Takikawa, Y., and Kawagoe, R. (2000). Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiol. Rev.* 80, 953–978. doi: 10.1152/physrev.2000.80.3.953
- Howe, M. W., Tierney, P. L., Sandberg, S. G., Phillips, P. E., and Graybiel, A. M. (2013). Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature* 500, 575–579. doi: 10.1038/nature12475
- Humphreys, G. W. and Riddoch, M. J. (2000). “One more cup of coffee for the road: object-action assemblies, response blocking and response capture after frontal lobe damage,” in *Executive Control and the Frontal Lobe: Current Issues*, eds W. X. Schneider, A. M. Owen, and J. Duncan (Berlin: Springer), 81–93.
- Jai, Y. Y., and Frank, L. M. (2015). Hippocampal–cortical interaction in decision making. *Neurobiol. Learn. Mem.* 117, 34–41. doi: 10.1016/j.nlm.2014.02.002
- Jin, X., Tecuapetla, F., and Costa, R. M. (2014). Basal ganglia subcircuits distinctively encode the parsing and concatenation of action sequences. *Nat. Neurosci.* 17, 423–430. doi: 10.1038/nn.3632
- Koch, C. and Ullman, S. (1987). “Shifts in selective visual attention: towards the underlying neural circuitry,” in *Matters of Intelligence*, ed L. M. Vaina (Berlin: Springer), 115–141.
- Krauzlis, R. J., Bogadhi, A. R., Herman, J. P., and Bollimunta, A. (2017). Selective attention without a neocortex. *Cortex* 102, 161–175. doi: 10.1016/j.cortex.2017.08.026
- Krauzlis, R. J., Bollimunta, A., Arcizet, F., and Wang, L. (2014). Attention as an effect not a cause. *Trends Cogn. Sci.* 18, 457–464. doi: 10.1016/j.tics.2014.05.008
- Krauzlis, R. J., Lovejoy, L. P., and Zénon, A. (2013). Superior colliculus and visual spatial attention. *Annu. Rev. Neurosci.* 36, 165–182. doi: 10.1146/annurev-neuro-062012-170249
- LeDoux, J. E. (2014). Coming to terms with fear. *Proc. Natl. Acad. Sci. U.S.A.* 111, 2871–2878. doi: 10.1073/pnas.1400335111
- Li, S., Cullen, W. K., Anwyl, R., and Rowan, M. J. (2003). Dopamine-dependent facilitation of LTP induction in hippocampal CA1 by exposure to spatial novelty. *Nat. Neurosci.* 6, 526–531. doi: 10.1038/nn1049
- Lovejoy, L. P., and Krauzlis, R. J. (2010). Inactivation of primate superior colliculus impairs covert selection of signals for perceptual judgments. *Nat. Neurosci.* 13, 261–266. doi: 10.1038/nn.2470
- Minsky, M. (1988). *Society of Mind*. New York, NY: Simon and Schuster.
- Montague, P. R., Dayan, P., and Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive hebbian learning. *J. Neurosci.* 16, 1936–1947.
- Pelosi, A., and Guarino, D. (2015). A commentary on: “Cortical and thalamic excitation mediate the multiphasic responses of striatal cholinergic interneurons to motivationally salient stimuli.” *Front. Neural Circ.* 9:15. doi: 10.3389/fncir.2015.00015
- Ponzi, A., and Wickens, J. (2009). “Cell assemblies in large sparse inhibitory networks of biologically realistic spiking neurons,” in *Advances in Neural Information Processing Systems*, eds M. I. Jordan, Y. LeCun, and S. A. Solla (Boston, MA: MIT Press), 1273–1280.
- Sherman, S. M., and Guillery, R. W. (2001). *Exploring the Thalamus*, Vol. 312. Cambridge MA: Academic Press.
- Sommer, M. A., and Wurtz, R. H. (2004). What the brain stem tells the frontal cortex. II. role of the sc-md-fef pathway in corollary discharge. *J. Neurophysiol.* 91, 1403–1423. doi: 10.1152/jn.00740.2003
- Sommer, M. A., and Wurtz, R. H. (2006). Influence of the thalamus on spatial visual processing in frontal cortex. *Nature* 444, 374–377. doi: 10.1038/nature05279
- Tepper, J. M., and Bolam, J. P. (2004). Functional diversity and specificity of neostriatal interneurons. *Curr. Opin. Neurobiol.* 14, 685–692. doi: 10.1016/j.conb.2004.10.003
- Threlfell, S., Lalic, T., Platt, N. J., Jennings, K. A., Deisseroth, K., and Cragg, S. J. (2012). Striatal dopamine release is triggered by synchronized activity in cholinergic interneurons. *Neuron* 75, 58–64. doi: 10.1016/j.neuron.2012.04.038
- Treisman, A. M., and Gelade, G. (1980). A feature-integration theory of attention. *Cogn. Psychol.* 12, 97–136.
- Truccolo, W., Hochberg, L. R., and Donoghue, J. P. (2010). Collective dynamics in human and monkey sensorimotor cortex: predicting single neuron spikes. *Nat. Neurosci.* 13, 105–111. doi: 10.1038/nn.2455
- von der Malsburg, C. (1994). “The correlation theory of brain function,” in *Models of Neural Networks*, eds E. Domany, J. L. van Hemmen, and K. Schulten (New York, NY: Springer), 95–119.
- Voorn, P., Vanderschuren, L. J., Groenewegen, H. J., Robbins, T. W., and Pennartz, C. M. (2004). Putting a spin on the dorsal–ventral divide of the striatum. *Trends Neurosci.* 27, 468–474. doi: 10.1016/j.tins.2004.06.006
- Zenke, F., Agnes, E. J., and Gerstner, W. (2015). Diverse synaptic plasticity mechanisms orchestrated to form and retrieve memories in spiking neural networks. *Nat. Commun.* 6:6922. doi: 10.1038/ncomms7922
- Zénon, A., and Krauzlis, R. J. (2012). Attention deficits without cortical neuronal deficits. *Nature* 489, 434–437. doi: 10.1038/nature11497

**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 Guarino. 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.