YALE JOURNAL OF BIOLOGY AND MEDICINE 92 (2019), pp.121-125.

PERSPECTIVES

YALE JOURNAL OF BIOLOGY AND MEDICINE
YJBM
CELEBRATING 90 YEARS 1928–2018

Graded Memory: A Cognitive Category to Replace Spatial Sustained Attention and Working Memory

Camilo D. Libedinsky[a,b,c,*] and Pedro F. Fernandez[c]

[a]*Department of Psychology, NUS, Singapore, Singapore;* [b]*Singapore Institute for Neurotechnology, NUS, Singapore, Singapore;* [c]*Institute of Molecular and Cell Biology, A*STAR, Singapore, Singapore*

In this opinion article we challenge the commonly-held notion that visuospatial working memory and visuospatial sustained selective attention are two ontologically different cognitive categories. We start by discussing the general idea of cognitive categories, and then review some of the key behavioral and neural evidence both in favor of and against the separability of these processes. We then discuss a theoretical framework that could be useful for understanding the neural implementations of cognitive categories. We conclude that the evidence is insufficient to support the assumption that spatial working memory and spatial sustained attention are independent categories, and that further experimentation is necessary to determine the ontological independence of the two processes.

INTRODUCTION

Intelligent organisms are able to perform actions that are hard to interpret as simple stimulus-response associations. For example, a prey animal that spots a predator at a distance may freeze in place, and prepare for the detection of threatening movements, so as to initiate an escape behavior. By observing behaviors such as this one, we infer the existence of internal cognitive states which have been historically subdivided into distinct cognitive categories. In the example above, we may postulate that the prey animal's "preparation for the detection of threatening movements" involves a specific cognitive category: attention (*i.e.* the animal is paying attention to the predator). Similarly, if the predator is spotted, but then disappears behind a rock, the prey animal's behavior may be equivalent (freeze in place), but we may postulate the existence of a different cognitive category: short-term memory (*i.e.* the animal is remembering the former location of the predator, or where the predator is likely to re-emerge from). Importantly, these subdivisions are often assumed to reflect real categories, rather than just being a convenient operational subdivision.

Two such cognitive categories are voluntary endogenous visuospatial selective sustained attention and voluntary endogenous visuospatial working memory. The qualifiers voluntary, endogenous, visuospatial, and selective are used to disambiguate the meaning of the word

*To whom all correspondence should be addressed: Camilo D. Libedinsky, National University of Singapore, Faculty of Arts and Social Sciences, Block AS4, #03-39, 9 Arts Link, Singapore 117570; Tel: +65 9271 1190 (HP), Email: camilo@nus.edu.sg.

Keywords: Graded memory, visuospatial sustained attention, visuospatial working memory

Author Contributions: CDL and PFF conceptualized and wrote the manuscript.

attention. *Voluntary* is meant to distinguish it from involuntary attention (*i.e.* when a sound attracts your attention whether you want it or not). *Endogenous* is meant to distinguish it from exogenous attention (*i.e.* what ultimately triggers the onset of attention is an internal factor, such as a memory recollection, rather than an external factor, such as a stimulus). *Visuospatial* is meant to distinguish it from other types of attention, such as object or feature attention. *Selective* is meant to distinguish it from organism-level attentional states, such as high arousal states. *Sustained* is meant to distinguish it from transient attentional states (*i.e.* when a brief sound temporarily attracts your attention to a location). Normally, most of these qualifiers are not used on working memory. However, as discussed below, these also apply to working memory. Working memory can be *voluntary* or involuntary (*i.e.* do *not* remember these words: "pink elephant." You are now involuntarily remembering "pink elephant" using working memory). Working memories can also be *endogenous* or exogenous (same "pink elephant" as an example of an exogenous working memory). Working memories can also be *visuospatial* or non-visuospatial (*i.e.* memories of objects or features). *Selective* and *sustained* are the only qualifiers that do not apply to working memory. *Selective* because there are no organism-level states of working memory to distinguish it from. *Sustained* because shorter periods of memory are not called "working" memory (*i.e.* iconic memory). Henceforth we refer to them as attention and working memory, but the discussion is not meant to extend to other types of attentions and memories.

Attention and working memory are generally considered to be distinct from each other. Attention may be broadly defined as the cognitive state that allows organisms to volitionally select one or more spatial locations for prolonged preferential sensory processing, normally in the presence of an attended visual stimulus. An animal is presumed to be in an attentional state if its behavior satisfies the following criteria: stimuli that fall within the attended location are more likely to be detected and discriminated [1,2], and animals respond faster to these when engaged in a task that requires a speeded response [3,4]. On the other hand, working memory may be defined as the cognitive process that allows organisms to volitionally hold and manipulate a limited amount of visuospatial information, for a limited amount of time, in the absence of a visual stimulus. An animal is presumed to be in a working memory state if its behavior satisfies the following criteria: locations that are no longer cued by the presence of a stimulus may be recalled at a later time (normally a few seconds), and this memory can be modified if required.

Here, we challenge the assumption that attention and working memory are different cognitive categories. We propose that they may be better understood as a single cognitive category, which we term *graded memory*. Graded memory refers to the process that maintains information, both in presence or absence of stimuli, and which can be used to recall and manipulate the information as well as improve perceptual and motor processing. It is graded, because it can exist in one of many possible levels, and it is a memory because it can be maintained in the absence of stimuli. An animal can be presumed to be in a state of graded memory if its behavior satisfies the criteria of both attention and working memory. We propose that states that are normally referred to as attentional are those with a high level of graded memory. And states that are normally referred to as maintaining working memory are those with some level of graded memory, above a certain threshold. This proposal is based on behavioral, neurophysiological, and theoretical considerations. This line of inquiry is one on *cognitive ontologies*: the discussion of how to define cognitive categories and how these interact [5-8]. Here, we narrowly focus on the question of the existence of attention and working memory, or alternatively, of the existence of graded memory. Correct definition of cognitive categories has strong implications for cognitive neuroscience and psychology, but also for the identification and treatment of disorders of the nervous system.

KEY BEHAVIORAL EVIDENCE

Consider the following situation. You sit in front of a computer monitor, with nothing on display except for the appearance of a small white spot in the center. You receive an instruction to fixate on the spot and press a button on the keyboard when a small white cross appears anywhere on the screen. You are told that there is a 90 percent chance that the cross will appear on the right side of the screen. As you sit, fixating on the spot, waiting for the white cross to appear, what cognitive state are you in? According to our prior definitions you would be in an attentional state, since you would be faster at responding, and more likely to detect the cross when it appears on the right side of the screen, compared to the left. In addition, according to our definitions, you would also be in a working memory state, since the right side of the screen, which is not currently being cued by a stimulus, can be recalled for some time after the instruction. In this simple example, attention and working memory cannot be dissociated, since you cannot be instructed to pay attention to a location without implicitly instructing you to remember the location to pay attention to. Likewise, it may be hard, if not impossible, to remember a specific location without paying attention to it [9].

It could be argued that attention and working memory could be two separate, but intricately connected cognitive categories. Perhaps attention could play an important role in working memory rehearsal [10]. However, an alter-

native interpretation, and one which we consider more parsimonious than the alternatives, is that attention and working memory are one and the same cognitive category, such that what we call working memory refers to some level of graded memory in the absence of a stimulus, and what we call attention refers to a higher level of graded memory in the presence or absence of a stimulus.

Tasks where attention and working memory can be clearly dissociated are those in which the stimulus being attended to is continuously present; in these cases attention would be deployed in the absence of working memory. We believe that this clean separation is not a logical necessity, since the definition of a cognitive category should not depend on operational or historical factors, but rather attempt to reflect a real cognitive process. We argue that under these task conditions, graded memory would be engaged, in a similar way as it would be engaged when the attended stimulus is not present. Below we review some of the key behavioral studies that have attempted to understand the relationship between attention and working memory, and reinterpret them in the context of graded memory.

In dual-task interference experiments, participants perform a task that is temporally embedded within a second task, to assess how much they interfere with each other. The logic is that if neural "processing resources" are shared between the different cognitive categories, this would be reflected as a disruption of the performance in both tasks, whereas if these resources are independent, this would be reflected as a lack of interference between the simultaneous tasks. A number of studies have concluded that, while attention and working memory do interfere with each other, the interference is weak compared to that observed when the two simultaneous tasks belong to the same cognitive category (*i.e.* working-memory/working-memory or attention/attention) [11-14]. This evidence appeared to provide a strong argument against the existence of graded memory, and in favor of the separation of attention and working memory. However, more recent studies have challenged these conclusions, by conducting more controlled experiments, and suggested instead that attention and working memory interfere with each other as much as they interfere with themselves [15-17]. Therefore, these more recent and robustly controlled studies provide evidence against the existence of a separation between attention and working memory, and instead provide support for the existence of graded memory.

KEY NEURAL EVIDENCE

While the analysis of behavior is a powerful tool to study cognitive categories, it is possible that tightly linked cognitive processes, such as those of attention and working memory, are not dissociable at the behavioral level. Thus, a parallel approach to study cognitive categories, is to measure brain activity during task engagement, in order to find dissociable neural processes associated with the different cognitive categories. While functional imaging studies have suggested an overlap in brain mechanisms between attention and working memory [18,19], providing tentative support for the existence of graded memory, this method does not have sufficient resolution to exclude the possibility that intermixed within a specific brain region, separate populations of neurons subserve the different cognitive functions. Thus, in the following section we will discuss single-neuron studies in primates to assess the feasibility of the existence of graded memory compared to attention and working memory.

Prefrontal and parietal regions have been implicated in both attention and working memory [20-24]. When monkeys perform a dual-task that involves maintaining a memory while simultaneously paying attention to a stimulus, neurons in the lateral prefrontal cortex show decreased ability to represent task-relevant information [25], and their activity may relate to attention, working memory, or both simultaneously [26,27]. Along the same lines, when the activities of individual lateral prefrontal neurons are compared across two interleaved (non-simultaneous) tasks, one that requires working memory of a visual feature (direction of motion), and another one that requires passive perception of a moving visual stimulus (which presumably attracts attention automatically), some neurons show selectivity during the working memory period, passive perception period, or both [28]. Taken together, these results appear to provide evidence that attention and working memory involve differentiable neural substrates, thus providing an argument against the existence of graded memory, and in favor of the existence of attention and working memory. A shortcoming of using these studies as evidence in favor or against the existence of graded memory is that, unlike human behavioral and imaging experiments, to our knowledge, no single-neuron study has directly compared attention-memory dual-task interference with attention-attention and memory-memory dual-task interference. Similarly, to our knowledge, no single-neuron study has compared attention and working memory under identical stimulus conditions (*i.e.* when there is no stimulus present in both conditions, or when a stimulus is present in both conditions). Thus, with the existing evidence we cannot dissociate processing related to the different cognitive operations, from activity related to processing of the visual stimuli. Regardless of these considerations, these previous studies suggest the possible separability of neural mechanisms of attention and working memory. However, recent studies have highlighted the importance of interpreting neuronal activity within the context of dynamical neural systems, rather

than relying exclusively on the interpretation of single-neuron responses [29]. This is particularly important in non-sensory areas, where strong recurrent connectivity leads to highly dynamic population response profiles [30-32], with hard to interpret single-neuron selectivities [33,34]. Taking this into account, the existence of neurons that are selectively activated in attention tasks (with a stimulus present) compared to working memory tasks (without a stimulus present) [26-28] may not necessarily be interpreted as evidence of separate mechanisms for attention and working memory. Their existence could also be interpreted as the inevitable outcome of providing different inputs to a recurrent network that is performing one function. Thus, current research does not provide strong evidence either in support of or against the existence of graded memory. Further studies, as highlighted above, are necessary to specifically test this hypothesis.

CONCLUSION

In conclusion, we propose that the most parsimonious way to interpret the behavioral, neurobiological, and theoretical literature is that what we call visuospatial sustained attention and visuospatial working memory are in fact one cognitive category, which we refer to as graded memory. We have restricted this discussion to the visuospatial domain, but a similar argument could be extended to other types of cognitive categories.

Our analysis is in line with that of Ikkai and Curtis (2011) [18]. After providing neuroimaging evidence that activations in attention and working memory tasks are not distinguishable in prefrontal and parietal areas, they propose that persistent activity carries information that can be used generally to support a variety of cognitions, including attention and working memory. They propose that a general prioritized map of space in prefrontal and parietal areas could be used by downstream areas to fulfill specific behavioral demands. We believe that sustained activity may well be the neural implementation of graded memory, although other alternatives may exist [35-40].

Finally, we believe that the proposal of the existence of graded memory (and the non-existence of attention and working memory) is not mere semantic discussion with little substance. If in fact attention and working memory can be subsumed under graded memory, the false distinction between attention and working memory has led to a number of interpretational errors in the literature, where results are forced to fit into a false dichotomy in convoluted and inelegant ways. That said, it is also possible that our hypothesis is wrong, and attention and working memory are indeed real, separable, cognitive categories, or that neither attention, working memory or graded memory are real categories. Future experiments are required to address this question, and generative models of cognitive and perceptual processes could help address issues that are hard to assess experimentally. But with today's evidence, we conclude that assuming *a-priori* their independent existence is not warranted.

Acknowledgments: We thank Yen Shih-Cheng for comments and suggestions on an earlier version of this manuscript. This work was supported by a grant from the Ministry of Education Tier 2 Academic Research Fund to C.D.L.

REFERENCES

1. Bashinski HS, Bacharach VR. Enhancement of perceptual sensitivity as the result of selectively attending to spatial locations. Percept Psychophys. 1980;28(3):241–8.
2. Downing CJ. Expectancy and visual-spatial attention: effects on perceptual quality. J Exp Psychol Hum Percept Perform. 1988;14(2):188–202.
3. Posner MI. Orienting of attention. Q J Exp Psychol. 1980;32(1):3–25.
4. Mangun GR, Buck LA. Sustained visual-spatial attention produces costs and benefits in response time and evoked neural activity. Neuropsychologia. 1998 Mar;36(3):189–200.
5. Price CJ, Friston KJ. Functional ontologies for cognition: the systematic definition of structure and function. Cogn Neuropsychol. 2005;22(3):262–75.
6. Poldrack RA. Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. Neuron. 2011;72(5):692–7.
7. Poldrack RA, Yarkoni T. From Brain Maps to Cognitive Ontologies: Informatics and the Search for Mental Structure. Annu Rev Psychol. 2016;67:587–612.
8. Janssen A, Klein C, Slors M. What is a cognitive ontology, anyway? Philos Explor. 2017;20(2):123–8.
9. Awh E, Armstrong KM, Moore T. Visual and oculomotor selection: links, causes and implications for spatial attention. Trends Cogn Sci. 2006;10(3):124–30.
10. Awh E, Jonides J. Overlapping mechanisms of attention and spatial working memory. Trends Cogn Sci. 2001;5(3):119–26.
11. Allen R, McGeorge P, Pearson D, Milne A. Multiple-target tracking: A role for working memory? Q J Exp Psychol. 2006;59(6):1101–16.
12. Fougnie D, Marois R. Distinct capacity limits for attention and working memory: evidence from attentive tracking and visual working memory paradigms. Psychol Sci. 2006;17(6):526–34.
13. Zhang H, Xuan Y, Fu X, Pylyshyn ZW. Do objects in working memory compete with objects in perception? Vis Cogn. 2010;18(4):617–40.
14. Tas AC, Luck SJ, Hollingworth A. The relationship between visual attention and visual working memory encoding: a dissociation between covert and overt orienting. J Exp Psychol Hum Percept Perform. 2016;42(8):1121–38.
15. Feng J, Pratt J, Spence I. Attention and Visuospatial Working Memory Share the Same Processing Resources. Front Psychol. 2012;3(APR).
16. Won BY, Jiang YV. Spatial working memory interferes

with explicit, but not probabilistic cuing of spatial attention. J Exp Psychol Learn Mem Cogn. 2015;41(3):787–806.
17. Lapierre MD, Cropper SJ, Howe PD. Shared processing in multiple object tracking and visual working memory in the absence of response order and task order confounds. PLoS One. 2017;12(4):e0175736.
18. Ikkai A, Curtis CE. Common neural mechanisms supporting spatial working memory, attention and motor intention. Neuropsychologia. 2011;49(6):1428–34.
19. Silk TJ, Bellgrove MA, Wrafter P, Mattingley JB, Cunnington R. Spatial working memory and spatial attention rely on common neural processes in the intraparietal sulcus. Neuroimage. 2010;53(2):718–24.
20. Gnadt JW, Andersen RA. Memory related motor planning activity in posterior parietal cortex of macaque. Exp Brain Res. 1988;70(1):216–20.
21. Funahashi S, Bruce CJ, Goldman-Rakic PS. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. J Neurophysiol. 1989;61(2):331–49.
22. Corbetta M, Shulman GL. Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci. 2002;3(3):201–15.
23. Goldberg ME, Bisley J, Powell KD, Gottlieb J, Kusunoki M. The role of the lateral intraparietal area of the monkey in the generation of saccades and visuospatial attention. Ann N Y Acad Sci. 2002;956:205–15.
24. Armstrong KM, Chang MH, Moore T. Selection and maintenance of spatial information by frontal eye field neurons. J Neurosci. 2009;29(50):15621–9.
25. Watanabe K, Funahashi S. Neural mechanisms of dual-task interference and cognitive capacity limitation in the prefrontal cortex. Nat Neurosci. 2014;17(4):601–11.
26. Lebedev MA, Messinger A, Kralik JD, Wise SP. Representation of Attended Versus Remembered Locations in Prefrontal Cortex. PLoS Biol. 2004;2(11):e365.
27. Messinger A, Lebedev MA, Kralik JD, Wise SP. Multitasking of attention and memory functions in the primate prefrontal cortex. J Neurosci. 2009;29(17):5640–53.
28. Mendoza-Halliday D, Martinez-Trujillo JC. Neuronal population coding of perceived and memorized visual features in the lateral prefrontal cortex. Nat Commun. 2017;8:15471.
29. Driscoll LN, Golub MD, Sussillo D. Computation through Cortical Dynamics. Neuron. 2018;98(5):873–5.
30. Churchland MM, Cunningham JP, Kaufman MT, Foster JD, Nuyujukian P, Ryu SI, et al. Neural population dynamics during reaching. Nature. 2012;487(7405):51–6.
31. Mante V, Sussillo D, Shenoy KV, Newsome WT. Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature. 2013;503(7474):78–84.
32. Remington ED, Narain D, Hosseini EA, Jazayeri M. Flexible Sensorimotor Computations through Rapid Reconfiguration of Cortical Dynamics. Neuron. 2018;98(5):1005–19.
33. Rigotti M, Barak O, Warden MR, Wang XJ, Daw ND, Miller EK, et al. The importance of mixed selectivity in complex cognitive tasks. Nature. 2013;497(7451):585–90.
34. Parthasarathy A, Herikstad R, Bong JH, Medina FS, Libedinsky C, Yen SC. Mixed selectivity morphs population codes in prefrontal cortex. Nat Neurosci. 2017;20(12):1770–9.
35. Mongillo G, Barak O, Tsodyks M. Synaptic theory of working memory. Science. 2008;319(5869):1543–6.
36. Lewis-Peacock JA, Norman KA. Competition between items in working memory leads to forgetting. Nat Commun. 2014;5:5768.
37. Stokes MG. 'Activity-silent' working memory in prefrontal cortex: a dynamic coding framework. Trends Cogn Sci. 2015;19(7):394–405.
38. Lundqvist M, Herman P, Miller EK. Working Memory: Delay Activity, Yes! Persistent Activity? Maybe Not. J Neurosci. 2018;38(32):7013–9.
39. Constantinidis C, Funahashi S, Lee D, Murray JD, Qi XL, Wang M, et al. Persistent Spiking Activity Underlies Working Memory. J Neurosci. 2018;38(32):7020–8.
40. Masse NY, Yang GR, Song HF, Wang XJ, Freedman DJ. Circuit mechanisms for the maintenance and manipulation of information in working memory. bioRxiv. 2018; https://doi.org/10.1101/305714.