

Journal Club

Editor's Note: These short reviews of recent *JNeurosci* articles, written exclusively by students or postdoctoral fellows, summarize the important findings of the paper and provide additional insight and commentary. If the authors of the highlighted article have written a response to the Journal Club, the response can be found by viewing the Journal Club at www.jneurosci.org. For more information on the format, review process, and purpose of Journal Club articles, please see <http://jneurosci.org/content/preparing-manuscript#journalclub>.

On the Short-Lived Nature of Working Memory: Drift and Decay in a Population-coding model

Benjamin Cuthbert¹ and Dominic Standage^{1,2}

¹Centre for Neuroscience Studies, Queen's University, Kingston, Ontario K7L 3N6, Canada, and ²School of Psychology, University of Birmingham, Birmingham B15 2TT, United Kingdom

Review of Schneegans and Bays

The temporary storage and manipulation of information are known as working memory and are crucial to many cognitive abilities, including perception, problem-solving, and decision-making. Working memory research has a long and rich history, and a great deal of work has focused on characterizing and quantifying the critical limitations of the working memory system. This task has proven to be difficult; although working memory is described as both capacity-limited and short-lived, the specific nature of these constraints remains hotly debated (Lewandowsky et al., 2009; Luck and Vogel, 2013; Ma et al., 2014).

Data from behavioral experiments have been used to support various descriptions of working memory storage limitations, from suggestions of a fixed item limit of 7 ± 2 (Miller, 1956) to models positing the existence of 3 or 4 storage “slots” (Luck and Vogel, 1997; Cowan, 2001). Recently, visual working memory tasks requiring the report of a probed item from a continuous feature dimension have provided evidence that the precision

of item recall decreases as more items are stored (Wilken and Ma, 2004). These behavioral data can be described by models incorporating a fixed number of discrete item representations (Zhang and Luck, 2008) but are also accounted for by models where an unlimited number of representations can be maintained with variable precision (van den Berg et al., 2012).

The temporal limitations of working memory are also up for debate (Lewandowsky et al., 2009; Barrouillet et al., 2012). Despite the intuitive simplicity of representations that decay over time, controlling for various confounds during the delay period of memory tasks has proven difficult (Jonides et al., 2008). For example, some suggest that spontaneous decay of verbal information in working memory is counteracted via subvocal rehearsal of task stimuli (Baddeley et al., 1975) or by a refreshing process driven by internal attention (Barrouillet et al., 2004, 2012). To control for these hypothesized processes, some verbal working memory experiments involve a concurrent unrelated distractor task (e.g., repeating a nonsense phrase aloud or performing simple arithmetic), which typically reduces recall performance (Oberauer et al., 2016). While this effect may support the decay theories above, distractor processing itself may require representation in working memory and thus compromise otherwise robust

task representations through an interference process (Oberauer et al., 2012).

The impact of elapsed time on working memory for visual stimuli is also unclear. While some delayed recall studies (Zhang and Luck, 2009; Shipstead and Engle, 2012; Souza and Oberauer, 2015) and memory-guided saccade experiments (Funahashi et al., 1989; White et al., 1994; Starc et al., 2017) report decreased performance with longer delay periods, others report no effect (Vogel et al., 2001; Gorgoraptis et al., 2011). Although these experiments do not involve a distractor task, other forms of delay period interference have been proposed (Oberauer et al., 2016). For example, there is evidence that memory-guided saccadic responses are biased toward stimuli from previous trials, an effect that is known as proactive interference (Papadimitriou et al., 2015).

Due to differences in task design and equivocal results, the temporal limitations of working memory remain contentious. In addition, the impact of elapsed time is typically studied in isolation from storage limitations; a mechanistic account that addresses both of these limitations would therefore be a step forward for the field.

A framework recently introduced by Schneegans and Bays (2018) attempts to take such a step. The authors build on previous work proposing a population coding account of working memory storage (Bays, 2014). Their population coding (or

Received July 23, 2018; revised Oct. 1, 2018; accepted Oct. 5, 2018.

We thank Dr. Martin Paré and Dr. Gunnar Blohm for advice and support. The authors declare no competing financial interests.

Correspondence should be addressed to Benjamin Cuthbert, Centre for Neuroscience Studies, Queen's University, 18 Stuart Street, Kingston, Ontario K7L 3N6, Canada. E-mail: b.cuthbert@queensu.ca.

<https://doi.org/10.1523/JNEUROSCI.1877-18.2018>

Copyright © 2018 the authors 0270-6474/18/3810241-03\$15.00/0

“neural resource”) model suggests that features of items stored in working memory are represented by tuned populations of neurons, and that recall involves decoding the stored features. This model accounts for the impact of set size on recall precision by assuming that total population activity is normalized. As the number of representations within the population increases, the activity representing each stored feature decreases. The sampling and decoding of progressively noisier representations can thereby account for increased behavioral response variability.

The authors extend this framework to address the temporal limitations of working memory recall by comparing two hypotheses. The “decay” hypothesis posits that the rate of population activity representing an item continuously decreases over time (their Fig. 1A, top). Accordingly, the increased delay duration would have essentially the same impact as increased set size. Alternatively, the “drift” hypothesis posits that total population activity remains fixed, but drifts stochastically along a continuous feature dimension over time (their Fig. 1B, top). To test these hypotheses behaviorally, the authors took advantage of a less-studied feature of working memory performance: response time (RT).

In previous work, Schneegans and Bays (2016) reported that increasing set size in a spatial working memory task resulted in monotonic increases in RT and variability. Consistent with work describing working memory retrieval as a decision process (Pearson et al., 2014), they showed that the RT data could be fit by a model involving the accumulation of evidence toward a response threshold. In the current work, Schneegans and Bays (2018) suggest that evidence for a particular feature value takes the form of neural spiking activity. In this case, a greater number of spikes representing a feature would result in a faster accumulation of evidence to the response threshold and therefore a faster RT. Within this framework, population activity normalization explains the impact of set size on RT: at higher set sizes a probed feature would have fewer spikes representing it, so the accumulation of evidence to the response threshold would take longer.

The decay and drift hypotheses considered by Schneegans and Bays (2018) therefore make contrasting behavioral predictions about the impact of delay duration. Reduced population activity in the decay model predicts an increase in both response time and variability (their Fig.

1A, bottom), whereas normalized activity in the drift model predicts increased response variability without affecting RT (their Fig. 1B, bottom).

Behavioral data were collected from human participants performing a multiple-item variant of the classic oculomotor delayed-response task (Funahashi et al., 1989; Schneegans and Bays, 2018, their Fig. 2), where both set size (1, 2, or 4 colored items) and delay duration (0.5, 1, 2, or 4 s) were manipulated. Task performance replicated the significant increase in RT with set size shown previously (Schneegans and Bays, 2016), and the authors found no significant effect of delay duration on RT. This finding directly falsifies the decay hypothesis under the assumptions made here. Finally, the authors incorporated the two hypotheses into their population-coding model. Fits for all variants of the drift model were better able to capture the data than any version of the decay model. Together, these findings show that the decay model implemented here cannot account for imprecision in task performance, but they do not necessarily show that the drift hypothesis is the only mechanism able to account for their data.

As discussed above, the impact of time on working memory representations is controversial, and alternatives to temporal decay have often focused on the role of proactive interference (Papadimitriou et al., 2015). The latter posits that, when a long delay period is used, the temporal distance between a given trial and the previous trial is compressed, allowing information from the previous trial to interfere with the maintenance of stimuli during the current trial. Behavioral data from working memory experiments with both visual stimuli (Shipstead and Engle, 2012; Souza and Oberauer, 2015) and verbal stimuli (Oberauer and Lewandowsky, 2013) support this hypothesis, showing that decline in working memory recall performance with longer delay periods can be eliminated by fixing the ratio of intertrial to delay period duration. Schneegans and Bays (2018) suggest that such a process would be unlikely to account for their data, but they acknowledge that this hypothesis remains to be tested.

The population-coding model (Bays, 2014; Schneegans and Bays, 2018) is able to account for many aspects of behavior in working memory tasks but does not specify a neural mechanism for the maintenance of tuned activity over time. Persistent spiking activity during the delay period of working memory tasks is widely

regarded as a neural correlate of working memory maintenance (Zylberberg and Stowbridge, 2017), and was first observed in PFC (Fuster and Alexander, 1971; Kubota and Niki, 1971). Persistent spiking has since been identified in many cortical areas, including posterior parietal cortex (Gnadt and Andersen, 1988; Chafee and Goldman-Rakic, 1998) and sometimes visual cortex (Supèr et al., 2001; Dotson et al., 2018). Cortical microcircuit models known as attractor networks suggest that this activity can be sustained via recurrent excitation due to the long decay time constant of NMDA receptors (Brunel, 2000; Compte et al., 2000). In combination with broad feedback inhibition, attractor networks form stable activity patterns given a tuned input. Critically, this tuned activity persists after the input is removed (Compte et al., 2000).

An alternative to persistent spiking activity posits that working memory instead relies on short-term activity-dependent changes in synaptic strength (Barak and Tsodyks, 2014; Mongillo et al., 2017). This mechanism is known as short-term synaptic facilitation and is typically associated with elevated levels of presynaptic Ca^{2+} , which allow repeated stimuli to elicit successively greater postsynaptic responses (Jackman and Regehr, 2017). Several computational models incorporating short-term synaptic facilitation suggest that activity patterns representing previously encoded information can be regenerated by subsequent network activation (Barak and Tsodyks, 2014). Recently, synaptic facilitation has also been proposed as a mechanism that could support proactive interference in the memory-guided saccade task (Kilpatrick, 2018).

Schneegans and Bays (2018) explicitly support a sustained activity account of working memory. They note that noise in a subclass of attractor networks sometimes referred to as “bump attractors” can cause feature-encoding activity to move stochastically along the represented feature dimension. Significant drift is not a general feature of bump attractor models, however (Standage and Paré, 2018), and has been shown to decrease with increasing network size (Compte et al., 2000). Thus, it is unclear whether drift occurs in the brain.

A recent attempt to identify a neural correlate of drift was made by Wimmer et al. (2014), who analyzed recordings from the PFC of monkeys performing a single-item oculomotor delayed-response task with a 3 s delay. They found that variabil-

ity in the tuned activity of PFC neurons correlated with saccade endpoint variability. This finding is consistent with the hypothesis that drift underlies behavioral imprecision, but evidence of a temporal correlation between tuning curve drift and behavioral imprecision is still lacking. The use of multiple delay durations and more than one memory item would allow for a more thorough investigation of the temporal dynamics proposed by Schneegans and Bays (2018).

While a causal link between the physiology of working memory and behavioral data must still be established, the updated population-coding model developed by Schneegans and Bays (2018) provides a framework that addresses two fundamental features of working memory storage. The authors formally implement both the temporal decay and stochastic drift of working memory representations within this framework, and find that temporal decay is unable to account for their experimental data. By making testable predictions for the impact of set size and delay duration on working memory representations, this research helps lay the groundwork for a more comprehensive understanding of the neural mechanisms underlying working memory.

References

- Baddeley AD, Thomson N, Buchanan M (1975) Word length and the structure of short-term memory. *J Verbal Learn Verbal Behav* 14:575–589. [CrossRef](#)
- Barak O, Tsodyks M (2014) Working models of working memory. *Curr Opin Neurobiol* 25:20–24. [CrossRef](#) [Medline](#)
- Barrouillet P, Bernardin S, Camos V (2004) Time constraints and resource sharing in adults' working memory spans. *J Exp Psychol Gen* 133:83–100. [CrossRef](#) [Medline](#)
- Barrouillet P, De Paeppe A, Langerock N (2012) Time causes forgetting from working memory. *Psychonom Bull Rev* 19:87–92. [CrossRef](#) [Medline](#)
- Bays PM (2014) Noise in neural populations accounts for errors in working memory. *J Neurosci* 34:3632–3645. [CrossRef](#) [Medline](#)
- Brunel N (2000) Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *J Comput Neurosci* 8:183–208. [CrossRef](#) [Medline](#)
- Chafee MV, Goldman-Rakic PS (1998) Matching patterns of activity in primate prefrontal area 8a and parietal area 7ip neurons during a spatial working memory task. *J Neurophysiol* 79:2919–2940. [CrossRef](#) [Medline](#)
- Compte A, Brunel N, Goldman-Rakic PS, Wang XJ (2000) Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb Cortex* 10:910–923. [CrossRef](#) [Medline](#)
- Cowan N (2001) The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behav Brain Sci* 24:87–185. [CrossRef](#) [Medline](#)
- Dotson NM, Hoffman SJ, Goodell B, Gray CM (2018) Feature-based visual short-term memory is widely distributed and hierarchically organized. *Neuron* 99:215–226.e4. [CrossRef](#) [Medline](#)
- Funahashi S, Bruce CJ, Goldman-Rakic PS (1989) Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J Neurophysiol* 61:331–349. [CrossRef](#) [Medline](#)
- Fuster JM, Alexander GE (1971) Neuron activity related to short-term memory. *Science* 173:652–654. [CrossRef](#) [Medline](#)
- Gnadt JW, Andersen RA (1988) Memory related motor planning activity in posterior parietal cortex of macaque. *Exp Brain Res* 70:216–220. [Medline](#)
- Gorgoraptis N, Catalao RF, Bays PM, Husain M (2011) Dynamic updating of working memory resources for visual objects. *J Neurosci* 31:8502–8511. [CrossRef](#) [Medline](#)
- Jackman SL, Regehr WG (2017) The mechanisms and functions of synaptic facilitation. *Neuron* 94:447–464. [CrossRef](#) [Medline](#)
- Jonides J, Lewis RL, Nee DE, Lustig CA, Berman MG, Moore KS (2008) The mind and brain of short-term memory. *Annu Rev Psychol* 59:193–224. [CrossRef](#) [Medline](#)
- Kilpatrick ZP (2018) Synaptic mechanisms of interference in working memory. *Sci Rep* 8:7879. [CrossRef](#) [Medline](#)
- Kubota K, Niki H (1971) Prefrontal cortical unit activity and delayed alternation performance in monkeys. *J Neurophysiol* 34:337–347. [CrossRef](#) [Medline](#)
- Lewandowsky S, Oberauer K, Brown GD (2009) No temporal decay in verbal short-term memory. *Trends Cogn Sci* 13:120–126. [CrossRef](#) [Medline](#)
- Luck SJ, Vogel EK (1997) The capacity of visual working memory for features and conjunctions. *Nature* 390:279–281. [CrossRef](#) [Medline](#)
- Luck SJ, Vogel EK (2013) Visual working memory capacity: from psychophysics and neurobiology to individual differences. *Trends Cogn Sci* 17:391–400. [CrossRef](#) [Medline](#)
- Ma WJ, Husain M, Bays PM (2014) Changing concepts of working memory. *Nat Neurosci* 17:347–356. [CrossRef](#) [Medline](#)
- Miller GA (1956) The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol Rev* 63:81–97. [CrossRef](#) [Medline](#)
- Mongillo G, Rumpel S, Loewenstein Y (2017) Intrinsic volatility of synaptic connections: a challenge to the synaptic trace theory of memory. *Curr Opin Neurobiol* 46:7–13. [CrossRef](#) [Medline](#)
- Oberauer K, Lewandowsky S (2013) Evidence against decay in verbal working memory. *J Exp Psychol Gen* 142:380–411. [CrossRef](#) [Medline](#)
- Oberauer K, Lewandowsky S, Farrell S, Jarrold C, Greaves M (2012) Modeling working memory: an interference model of complex span. *Psychonom Bull Rev* 19:779–819. [CrossRef](#) [Medline](#)
- Oberauer K, Farrell S, Jarrold C, Lewandowsky S (2016) What limits working memory capacity? *Psychol Bull* 142:758–799. [CrossRef](#) [Medline](#)
- Papadimitriou C, Ferdoash A, Snyder LH (2015) Ghosts in the machine: memory interference from the previous trial. *J Neurophysiol* 113:567–577. [CrossRef](#) [Medline](#)
- Pearson B, Raskevicius J, Bays PM, Pertzov Y, Husain M (2014) Working memory retrieval as a decision process. *J Vis* 14:2. [CrossRef](#) [Medline](#)
- Schneegans S, Bays PM (2016) No fixed item limit in visuospatial working memory. *Cortex* 83:181–193. [CrossRef](#) [Medline](#)
- Schneegans S, Bays PM (2018) Drift in neural population activity causes working memory to deteriorate over time. *J Neurosci* 38:4859–4869. [CrossRef](#) [Medline](#)
- Shipstead Z, Engle RW (2012) Interference within the focus of attention: working memory tasks reflect more than temporary maintenance. *J Exp Psychol* 39:277. [CrossRef](#) [Medline](#)
- Souza AS, Oberauer K (2015) Time-based forgetting in visual working memory reflects temporal distinctiveness, not decay. *Psychonom Bull Rev* 22:156–162. [CrossRef](#) [Medline](#)
- Standage D, Paré M (2018) Slot-like capacity and resource-like coding in a neural model of multiple-item working memory. *J Neurophysiol* 120:1945–1961. [CrossRef](#) [Medline](#)
- Starc M, Murray JD, Santamauro N, Savic A, Diehl C, Cho YT, Srihari V, Morgan PT, Krystal JH, Wang XJ, Repovs G, Anticevic A (2017) Schizophrenia is associated with a pattern of spatial working memory deficits consistent with cortical disinhibition. *Schizophr Res* 181:107–116. [CrossRef](#) [Medline](#)
- Supér H, Spekreijse H, Lamme VA (2001) A neural correlate of working memory in the monkey primary visual cortex. *Science* 293:120–124. [CrossRef](#) [Medline](#)
- van den Berg R, Shin H, Chou WC, George R, Ma WJ (2012) Variability in encoding precision accounts for visual short-term memory limitations. *Proc Natl Acad Sci U S A* 109:8780–8785. [CrossRef](#) [Medline](#)
- Vogel EK, Woodman GF, Luck SJ (2001) Storage of features, conjunctions, and objects in visual working memory. *J Exp Psychol* 27:92. [Medline](#)
- White JM, Sparks DL, Stanford TR (1994) Saccades to remembered target locations: an analysis of systematic and variable errors. *Vision Res* 34:79–92. [CrossRef](#) [Medline](#)
- Wilken P, Ma WJ (2004) A detection theory account of change detection. *J Vis* 4:1120–1135. [CrossRef](#) [Medline](#)
- Wimmer K, Nykamp DQ, Constantinidis C, Compte A (2014) Bump attractor dynamics in prefrontal cortex explains behavioral precision in spatial working memory. *Nat Neurosci* 17:431–439. [CrossRef](#) [Medline](#)
- Zhang W, Luck SJ (2008) Discrete fixed-resolution representations in visual working memory. *Nature* 453:233–235. [CrossRef](#) [Medline](#)
- Zhang W, Luck SJ (2009) Sudden death and gradual decay in visual working memory. *Psychol Sci* 20:423–428. [CrossRef](#) [Medline](#)
- Zylberberg J, Strowbridge BW (2017) Mechanisms of persistent activity in cortical circuits: possible neural substrates for working memory. *Annu Rev Neurosci* 40:603–627. [CrossRef](#) [Medline](#)