Experimental Brain Research (2019) 237:1899–1905
https://doi.org/10.1007/s00221-019-05533-y

RESEARCH ARTICLE

Diffusion modeling of interference and decay in auditory short-term memory

Tyler D. Bancroft[1,2] · **Tyler M. Ensor**[3] · **William E. Hockley**[2] · **Philip Servos**[2] · **Jeffery A. Jones**[2]

Received: 1 September 2018 / Accepted: 27 March 2019 / Published online: 16 May 2019

Abstract
Decay and interference are two leading proposals for the cause of forgetting from working and/or short-term memory, and mathematical models of both processes exist. In the present study, we apply a computational model to data from a simple short-term memory task and demonstrate that decay and interference can co-occur in the same experimental paradigm, and that neither decay nor interference alone can account for all cases of forgetting.

Keywords Decay · Interference · Short-term memory · Diffusion · Model · Auditory

Decay and interference co-exist in short-term memory for simple stimuli

Human memory is far from infallible. Memory failure in working memory and/or short-term memory (WM/STM) can easily be demonstrated experimentally and identified in everyday life. We use the term "memory failure" in a general sense, to avoid attributing an incorrect response in a memory task to any particular process or stage of memory—for example, to a failure of encoding, a failure of maintenance, or a failure of retrieval. Two proposed causes of memory failure in WM/STM are decay (the time-dependent degradation or corruption of information held in STM) and interference (the loss of information from STM due to stimuli presented or tasks performed during memory maintenance). Assorted theoretical views exist regarding which of these processes are responsible for memory failure in WM/STM, and under which circumstances (e.g., Barrouillet et al. 2004; Lewandowsky and Oberauer 2015; Oberauer and Lewandowsky 2013, 2014; Souza and Oberauer 2015). We will argue that both interference and decay can operate simultaneously in WM/STM.

Decay theories

Decay theories generally suggest that degradation or loss of mnemonic representations occurs during retention. On a gross level, it is easy to find reports of performance decreasing as a retention interval increases (e.g., Harris et al. 2001; Mercer and McKeown 2014; Sinclair and Burton 1996; Thorndike 1913). An influential model of decay, the time-based resource-sharing (TBRS) model, hypothesizes that mnemonic representations in WM decay while subjects are engaged in tasks that prevent rehearsal (Barrouillet et al. 2004). Rehearsal is not limited to articulatory rehearsal (e.g., Baddeley et al. 1984), but may include other processes, such as attentional refreshing (Camos et al. 2009). Impeding rehearsal of mnemonic content will result in the eventual loss of that content.

Interference theories

Interference in WM/STM has been extensively studied (e.g., Oberauer and Lin 2017); of particular relevance to the current study is intrusion-based interference, in which memory failure is due to the intrusion of irrelevant or unwanted information (a distractor) into memory. One suggested form of intrusion-based interference is overwriting,

---

✉ Tyler D. Bancroft
bancroft@stu.ca

[1] Department of Psychology, St. Thomas University, Fredericton, NB E3B 5G3, Canada

[2] Department of Psychology, Wilfrid Laurier University, Waterloo, ON N2L 3C5, Canada

[3] Department of Psychology, Memorial University of Newfoundland, St. John's, NL A1C 5S7, Canada

in which a distractor is encoded into memory, in the process displacing some or all of the previously existing memory representation(s) (e.g., Nairne 1990).

In a study by Lange and Oberauer (2005), subjects memorized lists of three phonologically dissimilar consonant–vowel–consonant (CVC) trigrams. After the study list, subjects read a second list of CVCs (distractors) out loud, and were then required to report the study list in the original order. Critically, the distractor list was built by selecting one of the study items and constructing the distractors such that each letter from the selected study item was present in one distractor (i.e., one distractor had the same first consonant as the study item, one had the same vowel, and one had the same second consonant). Overwriting theory suggests that greater interference would exist between the distractors and the targeted study item, relative to the other study items. Indeed, this was the case, with correct recall of the selected study item lower than correct recall of non-selected study items, presumably due to intrusions from distractors. Lange and Oberauer (2005) reported similar effects in subsequent experiments that used words as stimuli and phonemes as features, and Oberauer and Lange (2008) reported a similar experiment, in which a study list word suffered a similar degree of interference whether the distractor words were similar to the study word, or whether all the phonemes in the study list word were repeated in the distractor list.

Another form of intrusion-based interference does not require the loss of previous memory representations, but assumes that distractors are added to memory in a manner that makes it difficult or impossible to retrieve only the target, or to differentiate between the target and distractor. This does not exclude overwriting (i.e., the destruction or loss of existing information in memory), but also allows the case in which the distractor is encoded into STM without destroying previous representations, and performance reductions are due to difficulty in retrieving only the desired representation(s). It also includes the case where the distractor and previous representations are combined or merged (Overwriting, in this paradigm, is a “special case” of the concept of intrusion-based interference).

The clearest evidence for intrusion-based interference comes from recent studies on STM for very simple stimuli. These studies use simple stimuli (such as auditory pure tones or tactile vibrations), and the tasks require the encoding and retention of only a single stimulus property (such as auditory pitch or vibrational frequency). As such, the contents of memory can be represented as a single numerical value. We have referred to this as scalar STM (Bancroft et al. 2016). These experimental paradigms have two substantial advantages: first, it is relatively straightforward to construct models of what we believe to be held in memory, and second, it has been demonstrated that the neural systems underlying scalar STM represent information in a decodable fashion—specifically, in the firing rates of neurons maintaining stimulus representations (e.g., Romo et al. 1999).

Several studies by Bancroft et al. have produced evidence for intrusion-based interference in both tactile and auditory STM. Bancroft and Servos (2011) presented subjects with a to-be-remembered tactile vibration (the target) and a comparison vibration (the probe), and asked subjects to report whether the target and probe were the same vibrational frequency or different frequencies. A brief additional vibration (the distractor) was presented during the delay period between the target and probe presentations. Subjects were not required to encode the distractor. The frequency of the distractor was either the same as the target frequency or different. Critically, this meant that when the target and probe were of different frequencies, and the target and distractor were also of different frequencies, the distractor was shifted either toward the probe frequency, or away from the probe frequency. Subjects made more correct responses in the different probe/different (away-shift) distractor condition than in the different probe/different (toward-shift) distractor condition (i.e., more “different” responses). Bancroft et al. (2011) reported similar results. These results were interpreted as evidence that the probe was being compared against the distractor (to some degree), consistent with the intrusion into memory of the distractor.

More recently, Bancroft et al. (2016) extended these interference results to the auditory domain. Instead of using tactile vibrations as stimuli, they used auditory pure tones, and asked subjects to compare the pitch of the stimuli. Subjects were presented with a to-be-remembered pure tone (the target) followed by a comparison tone (the probe), and were instructed to report whether the target and probe were of the same frequency or different frequencies. Similar to Bancroft and Servos’ (2011) study, a distractor was presented during the delay period between target and probe. As in previous studies, in the probe-different condition, subjects performed better when the pitch of the distractor was shifted away from the probe pitch than toward the probe pitch, consistent with the intrusion of distractors into memory.

An ongoing debate exists regarding the relative roles (or even the existence) of decay and interference in working memory (see above). The simple nature of the stimuli in scalar STM tasks makes these tasks suitable for studying these issues, particularly by integrating experimental research with computational modeling. Bancroft et al. previously applied a neurocomputational model to data from vibrotactile STM tasks (Bancroft et al. 2013; Bancroft et al. 2014b). In addition, Bancroft et al. (2012) applied a diffusion model to the results of Bancroft and Servos’ (2011) study. It is this model we use in the present study.

Model

The diffusion model adapted for Bancroft et al.'s (2012) study is a phenomenological model of the decay of a simple, scalar memory trace held in STM (Kinchla and Smyzer 1967). A memory trace representing the stimulus frequency held in memory varies over the retention interval as a random walk, after which it is compared against the probe stimulus by a subtractive process. This can be denoted by two equations:

$$m = f_{\mathrm{initial}} + z(p - n), \quad (1)$$

$$y = \left| m - f_{\mathrm{probe}} \right|, \quad (2)$$

where m is the value stored in memory, f_{initial} is the frequency of the target stimulus, z is the size of each positive or negative step (i.e., the change in the value of m at each step), p is the number of positive steps, n is the number of negative steps, f_{probe} is the frequency of the probe stimulus, and y is the perceived difference between the memory trace and the frequency of the probe. A same/different decision is made by comparing the value of y against a threshold θ, making a "different" decision if y exceeded θ, and a "same" decision otherwise.

Bancroft et al. (2012) used the diffusion model to compare different versions of the intrusion-based interference hypothesis. Using the above diffusion model, they constructed three simulations of Bancroft and Servos' (2011) task. In the first, the distractor intruded into STM, becoming integrated with the representation of the target stimulus (by a weighted averaging of the memory trace and the distractor frequency); in the second, the distractor was integrated with the subsequent probe stimulus (by a weighted averaging of the probe and distractor frequencies), and this combination was compared with the representation of the target stimulus held in memory; in the third, the distractor was ignored and not incorporated into memory or the decision-making process. The first simulation was the quantitatively best-fitting and also the only simulation to replicate Bancroft and Servos' (2011) empirical interference effect, consistent with the intrusion-based interference hypothesis.

In the present study, we apply the same model to experimental data from Bancroft et al.'s (2016) study of interference in auditory STM.

Simulation 1

Simulation 1 tested Bancroft et al.'s (2016) assertion that distractors can intrude into short-term memory for auditory pitch, by fitting the diffusion model to the means from their Experiment 1.

Experimental method

We briefly recap Bancroft et al.'s (2016) experimental method here. Subjects were presented with a series of three auditory pure tones: a 1000 ms to-be-remembered tone (the target), a 250 ms distractor tone, and a 1000 ms comparison tone (the probe). Subjects were instructed to report whether the target and probe were of the same frequency or different frequencies. The target and distractor were separated by a 900 ms delay period, and the distractor and probe were separated by a 350 ms delay period. The target and probe were either the same frequency or different (by 5 Hz), and could assume frequencies of either 510 or 515 Hz. The distractor was either the same as the target, or shifted 5 Hz from the target frequency. Critically, when the target and probe were of different frequencies, and the distractor was also different from the target frequency, the distractor frequency could be shifted toward that of the probe (e.g., target: 510 Hz, distractor: 515 Hz, probe: 515 Hz) or away from that of the probe (e.g., target: 510 Hz, distractor: 505 Hz, probe: 515 Hz).

Simulation method

The relative weighting of the distractor (w) was set to 0.2, proportional to the duration of the distractor relative to the total duration of the distractor and target. Upon the "presentation" of the distractor, the memory trace was modified to be a weighted average of the trace value and the distractor frequency, with weights 1-w and w, respectively. The random walk then proceeded for the remainder of the delay period. The stepsize z and decision threshold θ were allowed to vary as free parameters. Initial parameter search ranges were selected based on ad hoc pilot simulations. The search range and search interval for z were [0.05, 0.25] and 0.01, respectively, and for θ, were [0, 20] and 0.1, respectively. 500,000 trials/condition (half that in each of the away-shift and toward-shift conditions) were simulated for each value of z, and θ was fit across the search range for each simulated dataset. Simulations were coded in C++.

Table 1 Mean proportions of correct responses from Bancroft et al.'s (2016) Experiment 1 and from Simulations 1 and 2

	Experimental	Simulation 1	Simulation 2A	Simulation 2B
Same probe/same distractor	0.67	0.67	0.58	0.68
Same probe/different distractor	0.64	0.65	0.57	0.68
Different probe/same distractor	0.57	0.57	0.54	0.57
Different probe/Different (away) distractor	0.67	0.65	0.49	0.57
Different probe/different (toward) distractor	0.48	0.50	0.59	0.57
z		0.15	0.15	0.125
θ		4.4	4.3	4.3
SS		0.0011	0.0580	0.0194
r^2		0.98	0.19	0.31

Results and discussion

Values from the simulation and experimental results from Bancroft et al.'s (2016) Experiment 1 are presented in Table 1. The best-fitting parameter values were $z = 0.15$ and $\theta = 4.4$, with $\mathrm{SS} = 0.0011$, and $r^2 = 0.98$. It is clear that the model is able to account for virtually all the variance in the experimental data, and also qualitatively reproduces the interference effect reported by Bancroft et al. (i.e., better performance in the away-shift than the toward-shift condition). Further, the non-zero value for z does imply decay of the original mnemonic representation, in the sense that the accuracy of the representation decreases as a function of time. The performance of Bancroft et al.'s (2016) auditory STM experiment can be accurately replicated by assuming the combination of the intrusion of a distractor into memory, and the time-dependent decay of the original mnemonic representation.

Simulation 2

Simulation 2 simulates two contrasting control conditions: one in which the distractor is integrated with the probe (Simulation 2A), and one in which the distractor is not integrated into memory at all (Simulation 2B).

Simulation method

Simulation 2 uses the same general simulation method as Simulation 1, with the following changes. In Simulation 2A, the distractor is held in memory as a separate trace that does not interact with the target trace. At probe presentation, the distractor trace is combined with the probe frequency to form a probe trace, in a weight-averaged manner (with the weights w and 1-w, respectively). The weight of the distractor (w) was set to 0.2, consistent with Simulation 1. The distractor trace varied as a random walk with the same walk parameters as the target trace. In Simulation 2B, the distractor is not encoded into memory at all (e.g., $w = 0$), and the probe frequency is compared against the target trace.

Free parameters were the same as in Simulation 1 (θ, z); z was allowed to vary over the range [0.05, 0.25], with an interval of 0.025, and θ was allowed to vary over the range [0, 20], with an interval of 0.1. 200,000 trials/condition (half that in each of the away-shift and toward-shift conditions) were simulated for each value of z, and θ was fit across the search range for each simulated dataset. Simulation 2 was coded in Python.

Results and discussion

Results from Simulation 2 are presented in Table 1. The fit provided by the best-fitting parameters for Simulation 2A is far worse than that from Simulation 1 ($\mathrm{SS} = 0.058$ vs. 0.0011, respectively), and little of the variance in the experimental data is explained by the model in Simulation 2A ($r^2 = 0.19$). Similarly, Simulation 2B also produced a poorer fit to the experimental data ($\mathrm{SS} = 0.0194$), and the model used also explained much less of the variance in the experimental data ($r^2 = 0.31$) than did Simulation 1. This supports the conclusion that distractor stimuli are indeed intruding into STM and becoming integrated with (or displacing) the representation of the target stimulus being held in memory, rather than being integrated with the probe, or not being added to memory in any fashion.

General discussion

The present study offers an integrative view of the debate regarding decay versus interference. Experimental data from scalar STM tasks cannot be explained by only a decay account of memory failure. Rather, both decay and interference processes can operate in WM/STM. This, of course, does not mean that both processes operate on any given task. However, it is clear that it is not an either/or situation. Contrasting the case of Simulation 1 (in which the distractor is encoded into memory) against Simulation 2B (in which the distractor is not encoded), it is clear to see that experimental data are substantially better-fit by a model in which the distractor is encoded into memory. Further, the best-fitting value for the random walk stepsize in Simulation 1 was 0.15, suggesting a time-dependent degradation of the precision of the memory trace. We did not simulate the trivial case in which $z = 0$, as this leads to a situation in which there is no trial-by-trial variability in the memory trace. In this case, an appropriately chosen value for theta will give perfect performance in most conditions, inconsistent with empirical data. This suggests that time-dependent degradation of the memory trace is, in fact, taking place. Similarly, setting $w = 0$ (in Simulation 2B) also produces results inconsistent with experimental data. It appears that both decay and interference are necessary to explain the experimental data in question.

It is hardly a secret that memory failure tends to increase with the duration information is held in STM. The salient question is Why? Let us begin from the obvious position that information held in STM is represented by some neural state that is characteristic of the information. What mechanisms might be responsible for maintaining this neural state?

One well-supported answer is persistent neural activity. In the case of scalar STM, the salient remembered information (e.g., auditory pitch, vibrotactile frequency, etc.) has been found to be represented in the firing rates of neurons in frontal cortex during the delay periods of tasks, with some neurons' firing rates being positive monotonic functions of stimulus frequency, and others being negative monotonic functions of stimulus frequency (Romo et al. 1999; Romo and Salinas 2003). The precise neural mechanism by which these active states are maintained is unclear. Maintenance of a stable state in a network of neurons can be accomplished via a number of methods (e.g., Goldman 2009; Miller 2016; Renart et al. 2003; Wang 2001), but it is unclear from experimental data which (if any) of these are most appropriate in the case of scalar STM (In fact, no computational model of scalar STM has yet to be shown to adequately reproduce low-level neural activity patterns found experimentally; see Jun et al. 2010). Indeed, it is unclear whether recorded neural activity during scalar STM tasks actually reflects a stable representation, or whether scalar STM might rely on a neural representation that is unstable over long durations, but which is sufficiently stable over the short intervals used in experiments to support memory performance (Romo and Salinas 2003). Certainly, in other scalar STM tasks (such as the comparison of durations), we find not just a reduction in performance, but a systematic shift toward perceiving mnemonic representations of stimulus duration as being shorter, as the retention interval is lengthened.

Assuming that mnemonic representations in scalar STM equate to persistent neural activity, what does it mean that those mnemonic representations appear to change in a fashion similar to a diffusion process? If we assume that the stimulus frequency held in memory is a function of the net firing rate of the neurons involved in memory maintenance, then presumably a change in the stimulus representation is due to a change in firing rates. Again, without knowing how persistent neural activity is achieved in the neural systems underlying scalar STM, we are speculating to some degree—but the apparent drift in mnemonic representation during the delay period would seem to point to a corresponding drift in firing rates. Given that there are (presumably) thousands to millions of neurons involved in maintaining scalar information in STM, it seems likely that the average firing rate of the maintenance population of neurons may drift over time, due to fluctuations in the firing rates of individual neurons. Indeed, we have previously demonstrated that a classic STM finding (the subjective shortening effect) can be explained by assuming that a scalar STM store does not maintain a constant firing rate perfectly across the delay period, but rather that firing rates tend to drift (Bancroft et al. 2014a).

Some proposals have suggested that elevated frontal activity is not indicative of memory retention, but rather of other processes, possibly executive and/or attentional (e.g., Postle 2006). Such suggestions would seem to have less traction in the case of scalar STM, as the actual contents of memory can be decoded directly from neural firing rates in frontal cortex. This is also consistent with recent applications of decoding methods to neuroimaging data that have uncovered feature-specific information in prefrontal cortex during WM/STM storage (Ester et al. 2015). However, the decoding of memory contents from frontal cortex obviously does not imply that it is the sole location of memory storage, and stimulus representations might also exist in more posterior regions.

Our argument that scalar STM maintenance resembles a diffusion process is, obviously, dependent on the ability of the model to fit the data, but is also dependent on the specificity of that fit. A model with many free parameters could fit a dataset simply due to the flexibility of the model, rather than any correspondence between the structure of the model and the actual cognitive and/or neural processes involved. That is not the case in the present study. The model used in the present study has a simple structure and very few free parameters, which appear to assume similar values across a variety of experimental designs. Further, the model has been applied to scalar STM tasks in different sensory domains (Bancroft et al. 2012), suggesting that actual memory dynamics are being captured by the model. It would be fruitful to generate experimental datasets with many more experimental conditions, to allow the development and testing of more complex models.

An alternative model of decay, Barrouillet et al.'s (e.g., 2004) time-based resource-sharing (TBRS), proposes that information in memory decays when cognitive resources are occupied by tasks other than memory maintenance. As such, memory capacity is limited by the amount of time spent not engaged in rehearsal. The model predicts a negative relationship between cognitive load and WM capacity, consistent with a variety of experimental findings (e.g., Barrouillet et al. 2004, 2007), although not exclusively (e.g., Doherty and Logie 2016).

In scalar STM interference tasks, such as Bancroft et al.'s (2016) Experiment 1, the durations of the distractors are constant, and the distractors vary only in pitch. By Eq. 3, all experimental conditions would have the same cognitive load. As such, the TBRS model would predict no systematic difference in performance for different distractor conditions. This, obviously, is not consistent with empirical results. As such, we are left with the conclusion drawn earlier, namely, that memory performance in Bancroft et al.'s studies require an intrusion-based interference method to account for results.

The present study presents computational and experimental evidence that both decay and interference operate in WM/STM. It should be noted that the work presented here departs from standard span- and verbally-based WM/STM tasks. It is, therefore, not unreasonable to question the degree to which these findings inform general WM/STM theory, rather than the particular case of WM/STM for very simple forms of information. We suggest that the privilege enjoyed by span tasks (usually verbal or verbalizable information) as the "gold standard" for studying WM/STM may be due to their prolific use, rather than to an inherent superiority over other WM/STM tasks, and therefore, we might just as reasonably ask whether findings from "standard" tasks do indeed inform general WM/STM theory. The ability of the brain to store and manipulate information over short periods is hardly limited to verbal stimuli, but rather appears to encompass any information that can be processed perceptually, and is generally retained in relevant areas of representational cortex (e.g., Bettancourt and Xu 2016; Christophel et al. 2012, 2015; Christophel and Haynes 2014; Lee et al. 2013; Mitchell and Cusack 2008; Pratte and Tong 2014; Riggall and Postle 2012; Savini et al. 2012; Sligte et al. 2013; Sreenivasan et al. 2014a, b; Ven et al. 2012; see Bancroft et al. 2014a, b; D'Esposito and Postle 2015; Lee and Baker 2016; Postle 2006, 2015; Sreenivasan et al. 2014a, b, for reviews). And, as shown in the present work, nontraditional WM/STM experiments have produced support for simultaneous decay and interference in WM/STM. The present study offers several directions for future research. First, it would be beneficial to examine the details around criterion (threshold) selection in scalar STM tasks. In this simulation, we assumed a single threshold value was constant across the experiment, and we weighted all experimental conditions equally when calculating fit. Empirically, it is possible that subjects adjust their criterion over the course of an experiment (and in this case, the different/away and different/toward conditions would likely be weighted less than other conditions, as there were fewer trials in the different/away and different/toward conditions than the other conditions). Experimental research to gather data on criterion changes, integrated with computational modeling, would illuminate this issue. It would also be informative to use the present findings to help develop a more biologically detailed model of scalar STM storage that obeys the same diffusion dynamics that we have shown to fit experimental data well. Finally, extending diffusion modeling to other scalar STM domains (such as memory for stimulus duration) would be useful for examining the relationships between different scalar STM domains.

Acknowledgements WEH, PS, and JAJ received support from NSERC Discovery Grants. TDB and TME were supported by NSERC scholarships. TDB was supported by a start-up grant and a Major Research Grant from St. Thomas University.

References

Baddeley AD, Lewis V, Vallar G (1984) Exploring the articulatory loop. Q J Exp Psychol 36:233–252

Bancroft T, Servos P (2011) Distractor frequency influences performance in vibrotactile working memory. Exp Brain Res 208:529–532

Bancroft TD, Servos P, Hockley WE (2011) Mechanisms of interference in vibrotactile working memory. PLoS One 6:e22518

Bancroft TD, Hockley WE, Servos P (2012) Diffusion modeling of interference in vibrotactile working memory. NeuroReport 23:255–258

Bancroft TD, Hockley WE, Servos P (2013) Irrelevant sensory stimuli interfere with working memory storage: evidence from a computational model of prefrontal neurons. Cogn Affect Behav Neurosci 13:23–34

Bancroft TD, Hockley WE, Servos P (2014a) A shared short-term memory system for stimulus duration and stimulus frequency. Can J Exp Psychol 68:236–241

Bancroft TD, Hockley WE, Servos P (2014b) Does stimulus complexity determine whether working memory storage relies on prefrontal or sensory cortex? Atten Percep Psychophys 76:1954–1961

Bancroft TD, Jones JA, Ensor TM, Hockley WE, Servos P (2016) Overwriting and intrusion in short-term memory. Mem Cognit 33:435–443

Barrouillet P, Bernardin S, Camos V (2004) Time constraints and resource sharing in adults' working memory spans. J Exp Psychol Gen 133:83–100

Barrouillet P, Bernardin S, Portrat S, Vergauwe E, Camos V (2007) Time and cognitive load in working memory. J Exp Psychol Learn Mem Cogn 33:570–585

Bettencourt KC, Xu Y (2016) Decoding the content of visual short-term memory under distraction in occipital and parietal areas. Nat Neurosci 19:150–157

Camos V, Lagner P, Barrouillet P (2009) Two maintenance mechanisms of verbal information in working memory. J Mem Lang 61:457–469

Christophel TB, Haynes J-D (2014) Decoding complex flow-field patterns in visual working memory. NeuroImage 91:43–51

Christophel TB, Hebart MN, Haynes J-D (2012) Decoding the contents of visual short-term memory from human visual and parietal cortex. J Neurosci 32:12983–12989

Christophel TB, Cichy RM, Hebart MN, Haynes J-D (2015) Parietal and early visual cortices encode working memory content across mental transformations. NeuroImage 106:198–206

D'Esposito M, Postle BR (2015) The cognitive neuroscience of working memory. Annu Rev Psychol 66:115–142

Doherty JM, Logie RH (2016) Resource-sharing in multiple-component working memory. Mem Cogn 44:1157–1167

Ester EF, Sprague TC, Serences JT (2015) Parietal and frontal cortex encode stimulus-specific mnemonic representations during visual working memory. Neuron 87:1–13

Goldman MS (2009) Memory without feedback in a neural network. Neuron 61:621–634

Harris JA, Harris IM, Diamond ME (2001) The topography of tactile working memory. J Neurosci 21:8262–8269

Jun JK, Miller P, Hernández A, Zainos A, Lemus L, Brody CD, Romo R (2010) Heterogenous population coding of a short-term memory and decision task. J Neurosci 30:916–929

Kinchla RA, Smyzer F (1967) A diffusion model of perceptual memory. Percept Psychophys 2:219–229

Lange EB, Oberauer K (2005) Overwriting of phonemic features in serial recall. Memory 13:333–339

Lee S-H, Baker CI (2016) Multi-voxel decoding and the topography of maintained information during visual working memory. Front Syst Neurosci 10:2

Lee S-H, Kravitz DJ, Baker CI (2013) Goal-dependent dissociation of visual and prefrontal cortices during working memory. Nat Neurosci 16:997–999

Lewandowsky S, Oberauer K (2015) Rehearsal in serial recall: an unworkable solution to the nonexistent problem of decay. Psychol Rev 122:674–699

Mercer T, McKeown D (2014) Decay uncovered in nonverbal short-term memory. Psychon Bull Rev 21:128–135

Miller P (2016) Dynamical systems, attractors, and neural circuits. F1000Research 5:992

Mitchell DJ, Cusack R (2008) Flexible, capacity-limited activity of posterior parietal cortex in perceptual as well as visual short-term memory tasks. Cereb Cortex 18:1788–1798

Nairne JS (1990) A feature model of immediate memory. Mem Cognit 18:251–269

Oberauer K, Lange EB (2008) Interference in verbal working memory: distinguishing similarity-based confusion, feature overwriting, and feature migration. J Mem Lang 58:730–745

Oberauer K, Lewandowsky S (2013) Evidence against decay in verbal working memory. J Exp Psychol Gen 142:380–411

Oberauer K, Lewandowsky S (2014) Further evidence against decay in working memory. J Mem Lang 73:15–30

Oberauer K, Lin H-Y (2017) An interference model of visual working memory. Psychol Rev 124:21–59

Postle BR (2006) Working memory as an emergent property of the mind and brain. Neuroscience 139:23–38

Postle BR (2015) The cognitive neuroscience of visual short-term memory. Curr Opin Behav Sci 1:40–46

Pratte MS, Tong F (2014) Spatial specificity of working memory representations in the early visual cortex. J Vis 14:22

Renart A, Song P, Wang X-J (2003) Robust spatial working memory through homeostatic synaptic scaling in heterogeneous cortical networks. Neuron 38:473–485

Riggall AC, Postle BR (2012) The relationship between working memory storage and elevated activity as measured with functional magnetic resonance imaging. J Neurosci 32:12990–12998

Romo R, Salinas E (2003) Flutter discrimination: neural codes, perception, memory and decision making. Nat Rev Neurosci 4:203–218

Romo R, Brody CD, Hernández A, Lemus L (1999) Neuronal correlates of parametric working memory in the prefrontal cortex. Nature 399:470–473

Savini N, Brunetti M, Babiloni C, Ferretti A (2012) Working memory of somatosensory stimuli: an fMRI study. Int J Psychophysiol 86:220–228

Sinclair RJ, Burton H (1996) Discrimination of vibrotactile frequencies in a delayed pair comparison task. Percept Psychophys 58:680–692

Sligte IG, van Moorselaar D, Vandenbroucke ARE (2013) Decoding the contents of visual working memory: evidence for process-based and content-based working memory areas? J Neurosci 33:1293–1294

Souza AS, Oberauer K (2015) Time-based forgetting in visual working memory reflects temporal distinctiveness, not decay. Psychon Bull Rev 22:156–162

Sreenivasan KK, Curtis CE, D'Esposito M (2014a) Revisiting the role of persistent neural activity during working memory. Trends Cognit Sci 18:82–89

Sreenivasan KK, Vytlacil J, D'Esposito M (2014b) Distributed and dynamic storage of working memory stimulus information in extrastriate cortex. J Cognit Neurosci 26:1141–1153

Thorndike EL (1913) Educational psychology: the psychology of learning, vol 2. Teachers College, New York

van de Ven V, Jacobs C, Sack AT (2012) Topographic contribution of early visual cortex to short-term memory consolidation: a transcranial magnetic stimulation study. J Neurosci 32:4–11

Wang X-J (2001) Synaptic reverberation underlying mnemonic persistent activity. Trends Neurosci 24:455–463