
Trends in Cognitive Sciences
Trends Cogn. Sci. March 2009 Vol. 13 No. 3, pp. 101–144 ISSN 1364-6613
Focus on gaze and social attention
Mirror neurons and speech perception
The allocation of attention in reading
Forgetting in verbal short-term memory
Cell PRESS

No temporal decay in verbal short-term memory

Stephan Lewandowsky[1], Klaus Oberauer[2] and Gordon D.A. Brown[1,3]

[1] Schoool of Psychology, University of Western Australia, Crawley, WA 6009, Australia
[2] Department of Experimental Psychology, University of Bristol, 12a Priory Road, Bristol, BS8 1TU, UK
[3] Department of Psychology, University of Warwick, Coventry, CV4 7AL, UK

Many models of short-term memory (STM) ascribe an important role to temporal decay and forgetting because of the passage of time alone. We argue against decay as the primary form of forgetting from STM, and suggest that new experimental methodologies and recent models provide new perspectives on the old issue of the causes of forgetting. We show that several classic sources of evidence for time-based forgetting can be reinterpreted in terms of an interference-based view, and that new experiments provide compelling evidence against decay. We conclude that progress requires moving beyond demonstrations of qualitative effects and focusing instead on testing quantitative predictions of models.

Crucial but fragile: memory for the short term and cognition

Cognition requires maintenance of, and access to, task-relevant information in memory. For example, mental arithmetic would be impossible without a short-term or working memory for intermediate sums. (Here, we do not differentiate between 'short-term memory' and 'working memory' because the causes of forgetting are relevant to both.) Indeed, working-memory capacity strongly predicts many cognitive functions (e.g. Ref. [1]). The capacity and duration of short-term memory (STM) are nonetheless strikingly limited: people rarely remember more than a handful of items and sometimes forgetting is complete after a second or two [2]. What explains these limitations?

Intuition might indicate that memories simply weaken with time. Two conceptions of such temporal forgetting have been embodied in contemporary models. One is that memory inexorably decays over time (e.g. Refs [3,4]), and the other is that traces become less temporally distinct, and thus more difficult to retrieve as time passes (e.g. Ref. [5]). We focus here on the popular decay (see Glossary) hypothesis because – unlike distinctiveness – it invariably predicts forgetting because of the passage of absolute time. Distinctiveness models are addressed in Box 1. Opposing a role of time *per se*, other theorists have suggested that forgetting results from subsequent events that interfering (e.g. Refs [6,7]).

There has been much renewed interest in this long-standing controversy, and new methodologies have emerged whereas traditional evidence such as the 'word-length effect' (Box 2) have been reinterpreted. We argue that this recent methodological progress, combined with rigorous quantitative modelling that reveals the differing predictions of theories, enables rejection of decay as the primary cause of short-term forgetting.

Adjudicating between decay and interference

What evidence could differentiate between time-based and interference-induced forgetting? This issue is non-trivial for two reasons. First, it is *a priori* unclear how much forgetting would be expected by the two views. If decay causes forgetting, how much forgetting is expected? Suppose recall declines from 0.80 to 0.79 after a 1 minute delay; is this evidence for decay? What about a decline from 0.80 to 0.50? Interpretation of relevant data is

Glossary

Articulatory suppression: the overt repeated articulation of irrelevant material (e.g. repeating a word aloud while studying or recalling a list, or both).

Complex-span task: the complex-span task is central to working-memory research. On each trial, a distractor (or 'processing') task alternates with to-be-remembered items. A typical sequence might be 2 + 3 = 5?, A, 5 + 1 = 7?, B, . . ., where the equations have to be judged for correctness and the letters must be memorized for immediate serial recall after completion of the sequence.

Decay: refers to the notion that memory fades over time without an additional identifiable causal agent. The decay notion must assume a compensatory process such as rehearsal whenever forgetting over time is absent.

Interference: on the one hand, interference is a theoretical notion that refers to a hypothetical process by which memory is impaired by interfering representations or cognitive operations. Several variants of interference theories await adjudication (Box 1). On the other hand, interference can be a descriptive term that refers to the deleterious effect of irrelevant stimuli on recall of the memoranda. A common distinction is made between two classes of interference:

-*Proactive interference:* refers to the effect of earlier events on recall of the memoranda. For example, forgetting in the Brown-Peterson paradigm is more pronounced on later trials, indicating that proactive interference from early trials renders memories more fragile.

-*Retroactive interference:* arises from events that follow the memoranda. For example, the effects of distractors during retrieval on the recall of an already-studied list represent retroactive interference.

Rehearsal: is typically understood as the overt or covert recitation of to-be-remembered material, which is thus preserved from forgetting. This articulatory rehearsal can be prevented by articulatory suppression. Recently, another mode of rehearsal has been postulated [3] and empirically supported [26,27] that relies on an attentional mechanism and that can demonstrably co-occur with overt verbalization (Box 2).

Temporal distinctiveness: temporal distinctiveness views (e.g. Ref. [5]) assume that memory performance is a function of time. However, unlike the decay notion, distinctiveness models attribute time-based forgetting to traces becoming temporally more crowded. In the same way that telegraph poles viewed from a moving train window become less distinguishable as they recede into the distance, events that recede into the past become more confusable and hence more difficult to retrieve.

Corresponding author: Lewandowsky, S. (lewan@psy.uwa.edu.au).

1364-6613/$ – see front matter © 2008 Elsevier Ltd. All rights reserved. doi:10.1016/j.tics.2008.12.003 Available online 14 February 2009

Box 1. Open questions

- If STM forgetting is best explained by interference rather than decay, how exactly does interference operate? Several theoretical options await exploration:

  Process-based interference. Processing activity can disrupt consolidation of items within the first 500 ms of presentation, presumably because both draw on an attentional bottleneck [30]. Engaging the bottleneck can also interfere with attentional refreshing [3].

  Interference by superposition. In models using distributed representations, such as SOB [6], items interfere with each other by being superimposed in a composite memory. The greater the number of items that are encoded, the more their representations are blurred.

  Interference by feature overwriting. With distributed item representations, interference can also operate through feature overwriting [7,31]. If items in STM share some features, those features could be deleted ('overwritten').Whereas superposition results in a general distortion of the memory trace, feature overwriting results in selective loss of shared features only.

  Interference by retrieval confusion. Interference could alternatively arise not from degradation of memory traces but because too many items have been associated with the available retrieval cues ('cue overload'; see Ref. [32]).
- If there is no decay, what is the role of rehearsal in STM? Rehearsal could be an epiphenomenon that is caused by memory contents without affecting them. Alternatively, rehearsal could counteract the effects of interference, for example by restoring overwritten features through redintegration [33].
- If there is no decay, what are the implications for temporal distinctiveness? Although distinctiveness often predicts time-based forgetting, and is thus challenged by the data reviewed here, relevant theories (e.g. SIMPLE [5]) postulate multiple attentionally weighted dimensions of distinctiveness, including one for ordinal list position. Therefore, SIMPLE predicts time-based forgetting only when people pay attention to time: independent evidence from manipulations of the temporal isolation of list items (e.g. Ref. [34]) indicates that people pay little attention to time in serial recall.

  Moreover, temporal-distinctiveness models assume temporal scale invariance: therefore, relative rather than absolute time determines forgetting, and increasing retention duration need not cause more forgetting if memory representations are also temporally re-scaled. Supporting evidence comes from the finding that forgetting on a final Brown-Peterson trial with a constant delay differs between groups as a function of the delay they experienced on prior trials [35]. These prior delays determine the time scales of people, relative to which the constant final delay was experienced as large or small.

Box 2. The word-length effect revisited

The word-length effect (WLE; Ref. [36]) is the finding that words that take longer to pronounce are sometimes remembered more poorly than short words. The fact that differences in pronunciation durations of only 150–200 ms per word result in poorer recall is assumed to arise because long words have more time to decay before they can be rehearsed or output.

We argue that the WLE provides no evidence for decay at all, for two reasons [37]. First, even when consideration is restricted to a purely duration-based WLE involving words of equal syllabic complexity but differing pronunciation durations (e.g. 'platoon' versus 'racket'), the WLE represents a correlation between two measures – articulation duration and memory. Articulation duration, in turn, is correlated with many other features that influence the memorability of a word. Notwithstanding commendably thorough attempts to the contrary [38], it is impossible in principle to know, let alone control, all of these correlated features. This lack of control opens the door for alternative explanations not involving decay. Accordingly, several studies have found the WLE to be limited to particular stimuli (e.g. Ref. [39]), and the only study that manipulated length of (pseudo-) words experimentally found no effect of duration [40].

Second, it is unclear whether a decay model even predicts the WLE. The WLE is predicted by the conjunction of two assumptions: (i) traces decay quickly and (ii) longer articulation duration implies longer postponement of rehearsal. The second assumption is plausible if articulatory rehearsal is the only means to refresh verbal memory traces. Recent research, however, has revealed a second, attentional process by which memory traces can be refreshed. Attentional refreshing is dissociated from articulatory rehearsal [27], and can proceed concurrently with overt articulation [26]. Specifically, memory has been found to improve when distractor sentences were read aloud more slowly (and continuously): the opposite would have been expected on the basis of decay, and the continuous overt articulation would have prevented conventional rehearsal [26]. It follows that some other restoration process must exist that is beneficial to memory and that can occur during overt articulation; this process must be part of any viable decay model. If memory traces can be 'attentionally' refreshed concurrently with articulation, there is no need to assume that articulation of longer words postpones refreshing for a longer time than articulation of short words. Thus, there is no compelling reason for a decay model to predict a WLE.

impossible without obtaining quantitative predictions from competing models; some representative predictions are shown in Figure 1. For forgetting data to be theoretically conclusive, they must be compared to such quantitative benchmarks.

Second, auxiliary processes can be invoked. For example, if forgetting is unexpectedly absent, decay theorists can assume the presence of surreptitious rehearsal (e.g. Ref. [8]). Empirical tests must therefore control rehearsal, for example by requiring overt articulation (i.e. repeating a word aloud) or by adding a demanding task such as signal detection (Box 3). Conversely, an interference view can account for unexpected forgetting by postulating that activity during retention interfered with memory (Box 3). Empirical tests must therefore not only control rehearsal but also take care that retention intervals are kept free of interference. These goals conflict: to disable rehearsal, there must be some cognitive activity, but this activity could also create interference. Two methodologies have recently emerged that satisfy both goals.

Berman and colleagues [9] exploited the fact that, in short-term recognition, negative probes that were on the study list for the preceding trial are rejected more slowly than novel lures. This effect reflects the persistence of memory for the preceding list, even though it is now irrelevant and retention or rehearsal of those items is counter-productive. Berman *et al.* [9] found that the effect diminished only negligibly when the inter-trial interval was increased from 0.3 to 10 s, but was eliminated by insertion of a single intervening study-test trial of equal (10 s) duration. Thus, information that is no longer relevant lingers undiminished over time unless cleared by intervening cognitive events.

Lewandowsky and colleagues [10] manipulated retention time while blocking rehearsal by articulatory suppression during recall. Participants were trained to recall at different speeds (0.4, 0.8, and 1.6 s per item). Although recall of the last item was delayed by over 5 s at the slowest compared to the fastest speed, this added delay reduced performance by only 0.05 (on a 0 to 1 accuracy scale).

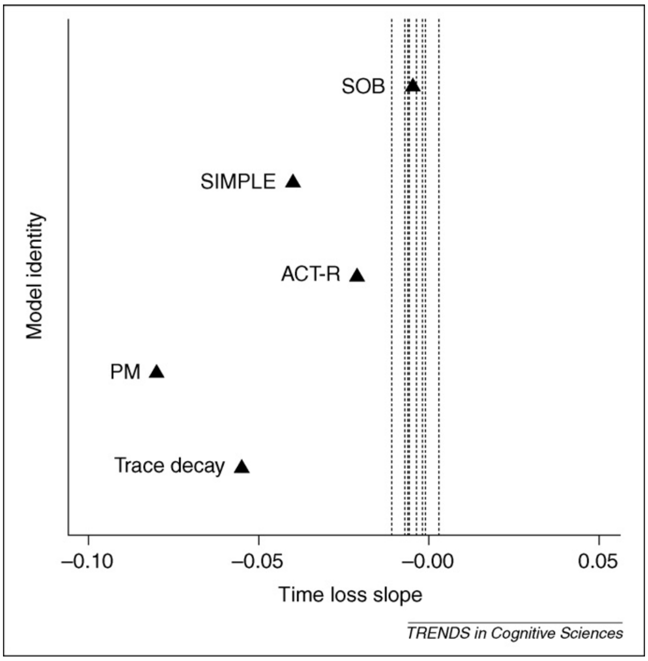

Figure 1. Observed time-loss slopes (TLS) from experiments in [10] and [12] are shown by vertical lines (one line for each study). TLS refers to the decline in absolute performance (i.e. proportion correct) per second additional delay. Linear TLS permit consideration of the extent of forgetting at each serial position separately, thus controlling output interference, by comparing performance between lists with long and short delays. The TLS values shown here are aggregated across serial positions. Rehearsal was prevented in all experiments by articulatory suppression, and in some studies additionally by an attention-demanding choice task. Black triangles show predicted TLS from one interference-based model (serial order in a box [SOB]; shown at the top) and four decay-based models (shown below the interference model). The Trace decay model [28] assumes a mean trace duration of 2 s, at which point the probability of recalling the whole list is 0.5. For a 6-item list, the probability of recalling each item under the assumption of independence must be $0.5^{1/6} = 0.89$, implying a loss of 0.11 in 2 s. Assuming linearity, TLS = 0.055. Predictions for ACT-R were obtained by running the model for the pilot experiment in Ref. [29], and varying the delay between presentation and recall. Predictions for the Primacy Model (PM), for SIMPLE and for SOB were computed with published parameter values. Adapted, with permission, from Ref. [12].

Cowan and colleagues [11] observed a similar result when children were either asked to recall at 'whatever speed seemed best' or 'as quickly as possible'. Recall times decreased from 0.82 s per item to 0.5 s per item (a speed-up of > 30%) but left recall essentially unchanged. Although rehearsal was not explicitly controlled, children were unlikely to have withheld their responses merely to rehearse when instructed to recall at a comfortable pace.

In a recent extension [12], yet another secondary task was added during retrieval to block possible 'attentional' forms of rehearsal (Box 2). In addition to overt articulation, participants performed a symbolic two-alternative choice task in between recalling list items. Increasing the number of articulations and choice responses from 1 to 4 substantially delayed recall (by up to 14 s for the last list item) but had only a negligible effect on memory (reducing accuracy by 0.005 per second additional delay). This observed forgetting rate constitutes a robust empirical target against which to examine theories (Figure 1). The same study [12] also introduced a varying number of distractors following list items during encoding, thus increasing retention duration for all items. This manipulation resulted in no additional forgetting whatsoever.

Box 3. Reitman's study revisited

In Reitman's classic experiments [41], people engaged in signal detection (a tone in white noise) during a 15 s retention interval. Participants who refrained from rehearsal were identified on the basis of their detection performance. In one experiment, those participants recalled some 34% less after 15 s than during immediate recall, leading Reitman to conclude that 'without surreptitious rehearsal, information in short-term memory decays' [41]. This conclusion seems premature for several reasons.

First, Reitman compared immediate recall without detection to a filled delay of 15 s, thus confounding the passage of time with the presence of a secondary task. We now know that a secondary task, however brief, severely impairs recall (e.g. Refs [12,42]) and that extending the duration of that task has little or no further detrimental effect [10,12]. Reitman's data are, thus, readily explained without recourse to decay, simply by acknowledging the known effects of a secondary task [42]. Although Reitman argued that the tones in her signal detection task bore little resemblance to the verbal memoranda, thus limiting similarity-based interference, there is now much evidence that interference can arise without any overt similarity, simply by the presence of another task [13,32]. To avoid this problem, an experiment would need to compare filled retention intervals of varying durations.

Second, each 15 s trial in Reitman's studies involved some seven signals. Each signal, whether detected or not, modulates the white noise, thus creating a 'changing-state' stream that is known to engender interference (e.g. Ref. [43]). To avoid this problem, analysis should consider only the occasional signal-free trials; however, their probability of occurrence was 0.00003.

In a recent conceptual replication of Reitman's study, both problems were resolved (Zwilling, C.E. Master's thesis, University of Missouri-Columbia, 2008): the duration of the (filled) retention interval was manipulated and analysis focused on signal-free 'catch' trials. Across three experiments, recall was found to be unaffected by increasing retention intervals, confirming that Reitman's original data were compromised.

Even putting aside methodological concerns, Reitman's results were inconsistent: in Experiment 1, subjects who were classified as non-rehearsers suffered a performance loss of around 0.02/s, whereas participants who met identical criteria in Experiment 2 lost less than 0.01/s (which is commensurate with recent results but below the rates predicted by decay models; see Figure 1 in main text). This discrepancy – a near tripling of forgetting between otherwise identical studies – is unsurprising, given that the results were based on six and four participants, respectively. (The confidence interval for absolute performance in Experiment 1 ranged from 0.49 to 0.83.)

To explain these results, proponents of inexorable decay would need to argue that some form of memory refreshing persisted, despite overt articulation blocking articulatory rehearsal and a symbolic choice task preventing attentional refreshing (Box 2). Decay proponents would also need to argue that people chose to rehearse a no-longer-relevant list despite this being counter-productive. We do not consider those arguments plausible. In addition, the recent work has established that the rate of forgetting in the absence of rehearsal or identifiable sources of interference is around 0.005/s. Current time-based models expect the rate of forgetting to be at least 0.02/s and normally considerably larger (Figure 1).

The time-based resource-sharing model: resurrection of decay?

A recent theory that combines decay with rehearsal is the time-based resource-sharing (TBRS) model of Barrouillet and colleagues [3]. Intriguingly, notwithstanding its reliance on decay, the TBRS need not predict increased

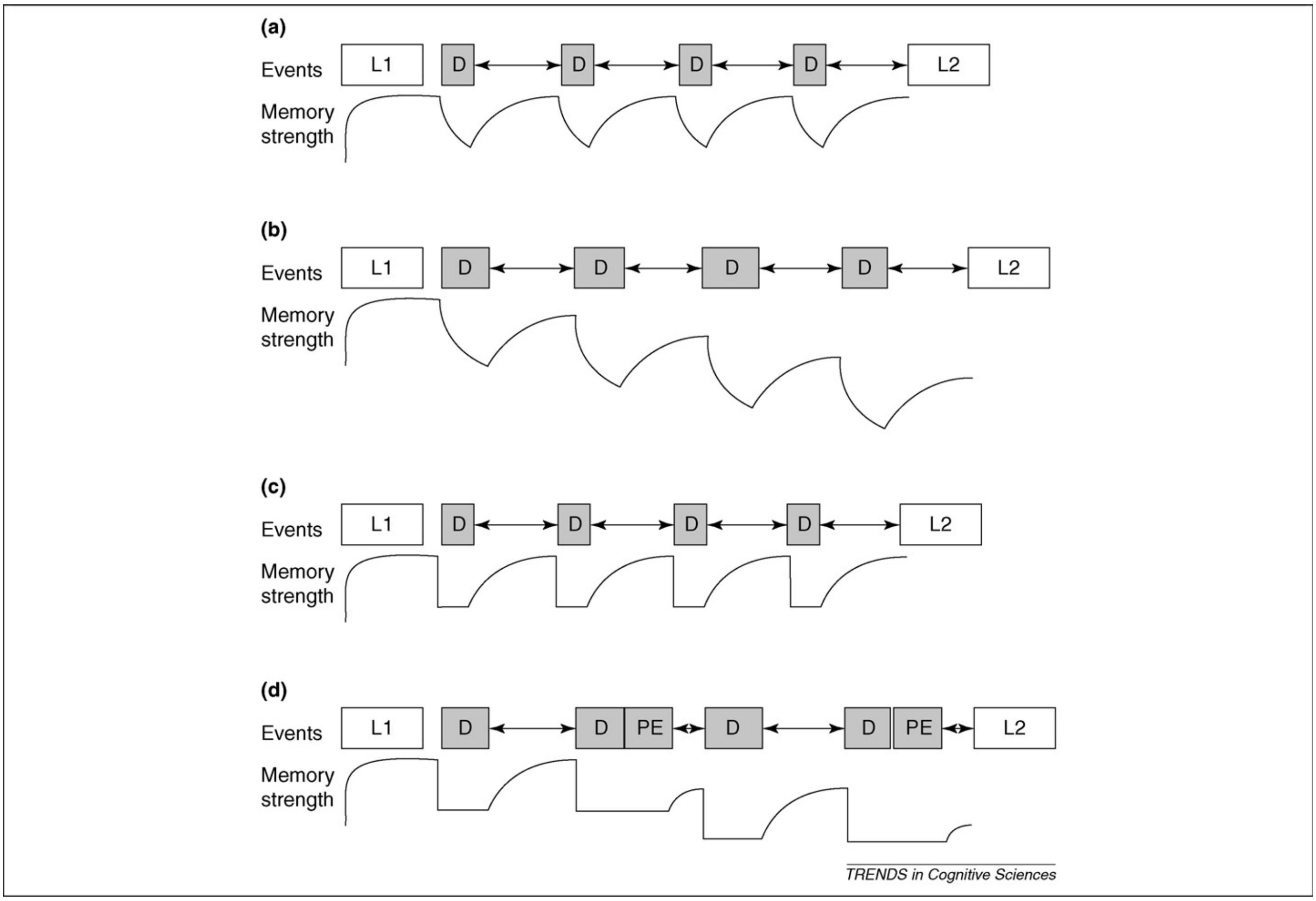

Figure 2. Time-line of events during one processing episode in between two memory items (L1 and L2) in a complex-span trial. **(a)** and **(c)** show easy, and **(b)** and **(d)** show difficult distractor tasks (shaded boxes, labelled D). Constant restoration intervals (arrows) follow each distractor. In each panel, the assumed evolution of memory strength of item L1 is shown below the time-line. (a) and (b) illustrate the assumption of the TBRS; (c) and (d) illustrate an alternative explanation not involving decay. (a) During short processing durations, memory decays; in the following restoration time, traces are attentionally refreshed, fully compensating decay. (b) During longer processing (more difficult distractors), memory decays more, such that the following refreshing episode cannot fully restore memory. (c) Distractors interfere with memory, reducing its strength independent of processing duration. In the following restoration time, memory traces are repaired. (d) Difficult decisions sometimes lead to errors, triggering post-error processing (PE) that occupies part of the restoration time. In the remaining time, memory cannot be fully restored. Portrat and colleagues [14] presented a study in which increased processing duration diminished recall. They interpreted the evidence within the decay notion [i.e. (a) and (b)], but a re-analysis reveals that the effect was not because of the passage of time but because of post-error processing [i.e. as in (c) and (d)].

forgetting if retention time is increased in a complex-span task. In TBRS, an attentional mechanism rapidly alternates between refreshing memory and performing the distractor task. Thus, forgetting is a function not of absolute time but of the proportion of time that is occupied by the distractor task ('cognitive load').

The TBRS's prediction that forgetting is a function only of cognitive load, rests on the assumption that the decay during each distractor activity is neutralized by refreshing in the 'restoration time' between that distractor and the next event (Figure 2a). Provided that refreshing completely reverses decay, adding more such decay-refreshing cycles does not add further forgetting; hence, overall duration of the retention interval does not matter. (In Figure 2a, more boxes (D) and arrows could be added without impairing performance on L1; for empirical confirmation see Refs [3,13].)

Complete reversal of decay by refreshing, however, can be assumed only with low cognitive loads (short boxes 'D' and long arrows in Figure 2a); at high loads (Figure 2b; boxes 'D' are longer relative to arrows than in Figure 2a), the TBRS predicts that refreshing is incomplete so that more forgetting should occur with additional distractors. In the experiments of Oberauer and Lewandowsky [12], participants were forced to process each distractor as quickly as possible (while also articulating continuously), and a new distractor appeared immediately after the response to the preceding one, thus much diminishing time for refreshing of memory traces (i.e. minimizing the arrows in Figure 2). The finding that a single distractor after each memory item led to substantial forgetting relative to no distractor confirms that this procedure created high cognitive load. Nevertheless, the number of distractors, and thus the duration of delay, had little or no effect on memory, contrary to what the TBRS must predict.

Another challenge to the TBRS arises from a recent study by Portrat and colleagues [14]. In their study, the duration of processing was manipulated via distractor difficulty while holding restoration time constant (i.e. Figure 2a versus Figure 2b). Memory was found to decline with additional processing time, thus seemingly implicating temporal decay. However, we discovered in a re-analysis of their data that this additional forgetting was entirely because of the increased error rate with the more difficult distractors. It is known that when people commit and then self-detect errors, the attendant post error processing occupies the attentional bottleneck [15]. In consequence, more difficult distractors encroach more upon the time available for memory restoration (Figure 2c,d). This alternative explanation was confirmed by our reanalysis: when both processing time and processing accuracy were entered as predictors of recall in a multilevel regression, only processing accuracy made a significant contribution, whereas processing time had no effect (parameter estimate 0.0002, corresponding to a 2/100th of a percent performance decline per second additional delay). The data of Portrat *et al.* [14] thus present, if anything, evidence against time-based forgetting.

Rapid forgetting by interference

What, if not decay, causes forgetting from STM? Recent evidence points to interference as the most important force. As noted earlier, a single intervening trial can accomplish forgetting that 10 s of delay does not [9]. Similarly, using serial recall with filled delays during retrieval, Lewandowsky and colleagues [16] showed that forgetting can be turned on and off by manipulating the to-be-articulated material. Unlike repetitions of the same word, articulation of continually changing material (e.g. 'super, table, house' rather than 'super, super super') led to considerable forgetting. It seems that forgetting occurs when the extent of interference associated with distracting material is enhanced.

Moreover, the same 'on-off' pattern also arises in the classic Brown-Peterson paradigm, in which a single extended period of distractor activity (such as reading and copying random letters [17]) follows presentation of the memoranda. When this period is over, recall proceeds without interruption. Massive and rapid forgetting can occur even after brief distractor periods (e.g. Ref. [18]). However, this forgetting is far from inevitable – it is much reduced or even eliminated when proactive or retroactive interference are minimized.

Concerning proactive interference, forgetting has long been known to be reduced or absent for the first trial in an experiment (e.g. Ref. [19]). When proactive interference is minimized by using a single experimental trial [17], the small residual forgetting can be explained by retroactive interference from the distractors in the delay period.

Evidence on the role of retroactive interference comes from two studies that minimized it in the same way as Lewandowsky *et al.* [16]; namely, by replacing a varied with a constant distractor activity in the retention interval. Very little forgetting was obtained for up to 15 s when the retention interval was filled by repeated articulation of the same word [8]. Using a similar distractor task, Phaf and Wolters [20] found small and usually non-significant forgetting over up to a minute.

Thus, contrary to common textbook summaries, the Brown-Peterson paradigm does not necessarily entail much if any time-based forgetting, and whether or not it does depends on the nature of the distractors or the presence of prior lists, as predicted on an interference view and contrary to the expectations of decay.

Box 4. Neuroscience of short-term forgetting

Can neuroscience inform the debate? Known or hypothesized basic neuronal mechanisms of memory seem compatible with either view of forgetting. Temporal decay could reflect neuronal firing rates that spontaneously decline over time [44], synaptic processes occurring over short time scales (e.g. Ref. [45]), or oscillatory mechanisms (see Ref. [46]). However, interference-based forgetting could also arise naturally from a variety of mechanisms, such as overloading the limited capacity of a neural network and overwriting of distributed neural representations (e.g. Ref. [47]).

As Wixted [32] notes, cognitive notions of interference and decay have developed separately from the focus of neuroscience on post-encoding consolidation failure as a source of memory loss. Memory traces that have not had time to consolidate are typically assumed to be vulnerable to interference (via degradation of the trace rather than because of cue overload at retrieval; see Box 1) by formation of new memory traces or other mental activity [32], but recent formal models have not provided sufficient details about neurobiological mechanisms to enable testable quantitative predictions.

Other neuropsychological data offer little direct support for decay. First, amnesic patients, who typically show near-complete forgetting after a minute or so, exhibit substantially less forgetting over periods of up to an hour when interference-minimizing conditions were created (e.g. when patients rest in a quiet room [48]), pointing to a crucial role for retroactive interference rather than the passage of time *per se* in amnesia.

Second, the involvement of different brain regions in short-term memory processes is well documented (e.g. Ref. [49]; see Ref. [47] for a review), and such research supports an interference-related role of specific brain regions. For example, frontal activation seems necessary to maintain representations in the focus of attention (e.g. Ref. [50]) and for resisting interference from distraction by irrelevant material during retention (e.g. Ref. [51]), in addition to resolving proactive interference [52]. Such findings sit naturally with an interference interpretation.

Moving forward

How might progress best be made in future? Despite pervasive agreement on interference rather than decay being 'the' cause of forgetting, at least in long-term memory [19,21], the concept of trace decay as a primary cause of short-term forgetting has proved resilient. We have argued that this reliance on decay is not justified by the data, including recent findings in neuroscience (Box 4).

Assuming an important role of interference in forgetting from STM seems inevitable, but might the slight performance loss ($\sim$0.005/s) that is observed in the absence of rehearsal and interference nonetheless mandate the inclusion of a time-based process in present or future models? Although this possibility cannot be excluded, to date the answer seems to be no. For example, Lewandowsky *et al.* [10] applied a temporal distinctiveness theory (SIMPLE; scale independent memory, perception and learning, see Ref. [5]) to forgetting data from individual participants and found that only one participant out of 11 was described significantly better by a variant of the theory that included time than by one that did not (in that version of SIMPLE, attention could be shifted between a temporal and a positional dimension, thus permitting a quantitative examination of the extent to which time affects performance). Likewise, Oberauer and Lewandowsky [12] showed that the data were best accommodated by a model that is entirely based on interference (SOB; e.g. Ref. [6]), and that decay models were unable to handle the results, even with optimized parameters. For example, the primacy model

[22] was unable to predict a rate of forgetting (TLS; time loss slope) below 0.035/s.

We suggest that models of short-term memory have now reached a level of sophistication in which it is no longer sufficient to look for qualitative effects. Instead, models and experimental work must proceed in tandem, by comparing precise measures of experimental effects with exact quantitative predictions (Figure 1).

We identified three ways in which models can contribute. First, they must be used to predict the exact amount of expected forgetting. We have shown that extant time-based models can be excluded because their quantitative predictions do not mirror the data.

Second, models can provide alternative explanations for phenomena that originally seemed to implicate decay. For example, Cowan and colleagues [23] argued for the presence of trace decay in a tone-comparison task, but Brown and colleagues [5] showed by simulation that the results could be explained without recourse to decay. Similarly, even before conceptual arguments against the interpretation of word-length effects were advanced (Box 2), computational models could explain the phenomena without appeal to decay or rehearsal (e.g. Refs [24,25]).

A third model-based analysis technique was illustrated earlier. If a model incorporates both time-based and interference-based forgetting, and the relative extent of each can be varied with a model parameter, then the model can be fitted to individual participant data and the relative contribution of each type of forgetting can be estimated from the corresponding parameter [10].

Further progress must await future models and data, but we suggest that at present there is little justification for reliance on decay as the only or even primary cause of forgetting from STM.

Acknowledgements

Preparation of this paper was facilitated by a Discovery Grant from the Australian Research Council and an Australian Professorial Fellowship to S.L., and a Linkage International Grant from the Australian Research Council to all three authors.

References

1 Oberauer, K. *et al.* (2007) Individual differences in working memory capacity and reasoning ability. In *Variation in Working Memory* (Conway, A.R.A. *et al.*, eds), pp. 49–75, Oxford University Press
2 Muter, P. (1980) Very rapid forgetting. *Mem. Cognit.* 8, 174–179
3 Barrouillet, P. *et al.* (2004) Time constraints and resource sharing in adults' working memory spans. *J. Exp. Psychol. Gen.* 133, 83–100
4 Burgess, N. and Hitch, G.J. (2006) A revised model of short-term memory and long-term learning of verbal sequences. *J. Mem. Lang.* 55, 627–652
5 Brown, G.D.A. *et al.* (2007) A temporal ratio model of memory. *Psychol. Rev.* 114, 539–576
6 Farrell, S. and Lewandowsky, S. (2002) An endogenous distributed model of ordering in serial recall. *Psychon. Bull. Rev.* 9, 59–79
7 Nairne, J.S. (1990) A feature model of immediate memory. *Mem. Cognit.* 18, 251–269
8 Vallar, G. and Baddeley, A.D. (1982) Short-term forgeting and the articulatory loop. *Q. J. Exp. Psychol.* 34, 53–60
9 Berman, M.G. In search of decay in verbal short-term memory. *J. Exp. Psychol. Learn. Mem. Cogn.* (In press)
10 Lewandowsky, S. *et al.* (2004) Time does not cause forgetting in short-term serial recall. *Psychon. Bull. Rev.* 11, 771–790
11 Cowan, N. *et al.* (2006) Rethinking speed theories of cognitive development: increasing the rate of recall without affecting accuracy. *Psychol. Sci.* 17, 67–73
12 Oberauer, K. and Lewandowsky, S. (2008) Forgetting in immediate serial recall: decay, temporal distinctiveness, or interference? *Psychol. Rev.* 115, 544–576
13 Barrouillet, P. *et al.* (2007) Time and cognitive load in working memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 33, 570–585
14 Portrat, S. *et al.* (2008) Time - related decay or interference-based forgetting in working memory? *J. Exp. Psychol. Learn. Mem. Cogn.* 34, 1561–1564
15 Jentzsch, I. and Dudschig, C. Why do we slow down after an error? Mechanisms underlying the effects of posterror slowing. *Q. J. Exp. Psychol.* (In press)
16 Lewandowsky, S. *et al.* (2008) Interference-based forgetting in verbal short-term memory. *J. Mem. Lang.* 59, 200–222
17 Baddeley, A.D. and Scott, D. (1971) Short-term forgetting in the absence of proactive interference. *Q. J. Exp. Psychol.* 23, 275–283
18 Peterson, L.R. and Peterson, M.J. (1959) Short-term retention of individual verbal items. *J. Exp. Psychol.* 58, 193–198
19 Keppel, G. and Underwood, B.J. (1962) Proactive inhibition in short-term retention of single items. *J. Verbal Learn. Verbal Behav.* 1, 153–161
20 Phaf, R.H. and Wolters, G. (1993) Attention shifts in maintenance rehearsal. *Am. J. Psychol.* 106, 353–382
21 McGeoch, J. (1932) Forgetting and the law of disuse. *Psychol. Rev.* 39, 352–370
22 Page, M.P.A. and Norris, D. (1998) The primacy model: a new model of immediate serial recall. *Psychol. Rev.* 105, 761–781
23 Cowan, N. *et al.* (1997) The role of absolute and relative amounts of time in forgetting within immediate memory: the case of tone pitch comparisons. *Psychon. Bull. Rev.* 4, 393–397
24 Brown, G.D.A. and Hulme, C. (1995) Modelling item length effects in memory span: no rehearsal needed? *J. Mem. Lang.* 34, 594–621
25 Lewandowsky, S. and Farrell, S. (2000) A redintegration account of the effects of speech rate, lexicality, and word frequency in immediate serial recall. *Psychol. Res.* 63, 163–173
26 Hudjetz, A. and Oberauer, K. (2007) The effects of processing time and processing rate on forgetting in working memory: testing four models of the complex span paradigm. *Mem. Cognit.* 35, 1675–1684
27 Raye, C.L. *et al.* (2007) Refreshing: a minimal executive function. *Cortex* 43, 135–145
28 Schweickert, R. and Boruff, B. (1986) Short-term memory capacity: magic number or magic spell? *J. Exp. Psychol. Learn. Mem. Cogn.* 12, 419–425
29 Anderson, J.R. and Matessa, M. (1997) A production system theory of serial memory. *Psychol. Rev.* 104, 728–748
30 Jolicoeur, P. and Dell'Acqua, R. (1998) The demonstration of short-term consolidation. *Cognit. Psychol.* 36, 138–202
31 Oberauer, K. and Kliegl, R. (2006) A formal model of capacity limits in working memory. *J. Mem. Lang.* 55, 601–626
32 Wixted, J.T. (2004) The psychology and neuroscience of forgetting. *Annu. Rev. Psychol.* 55, 235–269
33 Schweickert, R. (1993) A multinomial processing tree model for degradation and redintegration in immediate recall. *Mem. Cognit.* 21, 168–175
34 Lewandowsky, S. *et al.* (2008) When temporal isolation benefits memory for serial order. *J. Mem. Lang.* 58, 415–428
35 Turvey, M. *et al.* (1970) Proactive interference in short-term memory as a function of prior-item retention interval. *Q. J. Exp. Psychol.* 22, 142–147
36 Baddeley, A.D. *et al.* (1975) Word length and structure of short-term memory. *J. Verbal Learn. Verbal Behav.* 14, 575–589
37 Lewandowsky, S. and Oberauer, K. (2008) The word length effect provides no evidence for decay in short-term memory. *Psychon. Bull. Rev.* 15, 875–888
38 Mueller, S.T. *et al.* (2003) Theoretical implications of articulatory duration, phonological similarity, and phonological complexity in verbal working memory. *J. Exp. Psychol. Learn. Mem. Cogn.* 29, 1353–1380
39 Lovatt, P. *et al.* (2000) The word-length effect and disyllabic words. *Q. J. Exp. Psychol.* 53, 1–22
40 Service, E. (1998) The effect of word length on immediate serial recall depends on phonological complexity, not articulatory duration. *Q. J. Exp. Psychol.* 51, 283–304

41 Reitman, J.S. (1974) Without surreptitious rehearsal, information in short-term memory decays. *J. Verbal Learn. Verbal Behav.* 13, 365–377

42 Kane, M.J. and Engle, R.W. (2000) Working-memory capacity, proactive interference, and divided attention: limits on long-term memory retrieval. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 336–358

43 Jones, D.M. *et al.* (2004) The phonological store of working memory: is it phonological and is it a store? *J. Exp. Psychol. Learn. Mem. Cogn.* 30, 656–674

44 Fuster, J.M. (1973) Unit activity in prefrontal cortex during delayed response performance: neuronal correlates of transient memory. *J. Neurophysiol.* 36, 61–78

45 Zucker, R.S. and Regehr, W.G. (2002) Short-term synaptic plasticity. *Annu. Rev. Physiol.* 64, 355–405

46 Lustig, C. *et al.* (2005) Not "just" a coincidence: frontal-striatal interactions in working memory and interval timing. *Memory* 13, 441–448

47 Jonides, J. *et al.* (2008) The mind and brain of short-term memory. *Annu. Rev. Psychol.* 59, 193–224

48 Cowan, N. *et al.* (2004) Verbal recall in amnesiacs under conditions of diminished retroactive interference. *Brain* 127, 825–834

49 Henson, R.N.A. *et al.* (2000) Recoding, storage, rehearsal and grouping in verbal short-term memory: an fMRI study. *Neuropsychologia* 38, 426–440

50 Ranganath, C. and D'Esposito, M. (2005) Directing the mind's eye: prefrontal, inferior and medial temporal mechanisms for visual working memory. *Curr. Opin. Neurobiol.* 15, 175–182

51 Postle, B.R. (2006) Working memory as an emergent property of the mind and brain. *Neuroscience* 139, 23–38

52 Feredoes, E. *et al.* (2006) Direct evidence for a prefrontal contribution to the control of proactive interference in verbal working memory. *Proc. Natl. Acad. Sci. U. S. A.* 103, 19530–19534