

# Long term implicit and explicit memory for briefly studied words

Lee Averell (Lee.Averell@newcastle.edu.au)

School of Psychology, University of Newcastle.

Andrew Heathcote (Andrew.Heathcote@newcastle.edu.au)

School of Psychology, University of Newcastle.

## Abstract

Memories fade over time, but do they disappear altogether? The persistence of overlearned material has been regarded as evidence for permanent memories (Baird, 1984). However there seems a general consensus in the cognitive literature that briefly studied stimuli disappear from memory altogether (Wixted, 2004*ab*). We present evidence from implicit (stem completion) and explicit (stem cued recall) memory tasks showing well above chance performance 28 days after only one brief study event. Retention measured by both the implicit and explicit tasks was stable at the same level from seven days to 28 days. Our results question the consensus about the fate of memories for briefly studied stimuli.

Keywords: Forgetting; long term memory; recall; context.

## The Persistence of Long Term Memory

Whether or not stored memories are permanent is a subject of debate in both the cognitive and neuroscience communities. Here we investigate the issue of memory permanence and the way in which forgetting progresses over time in both explicit and implicit memory paradigms (Jacoby, Toth & Yonelinas, 1993; Squire, 1994). We first review previous pertinent findings and then report data from a stem-cued recall (explicit) and a stem completion (implicit) task testing memory at delays ranging from one minute to one month. The results are discussed with reference to novel context reinstatement techniques used in our experiment and their effects on mnemonic interference.

A number of studies have reported accurate explicit recall of information memorized years and even decades ago. Baird's (1984) seminal study showed that people could remember Spanish language words learned in a classroom for up to 50 years despite not using Spanish since the end of their high school education. He concluded that some memories are relatively permanent, fading only because of the detrimental effects of aging. More recently Squire (1989) found a similar pattern to Baird with memory for television programs from over 15 years prior to testing. Schmidt, Peck, Paas and van Breukelen (2000) showed that people could remember street names from suburbs they lived in as a child. Noice and Noice (2002) showed that actors could recall aspects of a Shakespearean script up to 28 years after playing the role. Results such as these support Checlie's (2006) statement that: "the inability ... to account for permanent retention is a serious failing" (p.36).

However, these studies of long term retention may be criticized on a number of grounds. They employed a cross sectional design that precludes investigation of the way in

which individual memory traces are forgotten. Further, given the meaningful nature of the material remembered in these studies it is possible that participants were reminded of and rehearsed the material in the study-test lag. It would, therefore, be desirable to examine retention in a longitudinal design of material that is less likely to be rehearsed.

The stimuli used in these studies were also initially very well learned. Hence, even if these studies demonstrate very long term retention, memory permanence may be restricted to such "overlearned" material. Baird (1984) suggested that participation in multiple spaced Spanish language courses in high school was a strong contributing factor to his results. Also, the effect of overlearning was evident in the differences in long term retention between higher achieving (A) students and less accomplished (B) students, with (A) students showing stronger retention in both the short and long term.

Evidence for very long term retention of stimuli studied only once has been reported in the implicit memory literature. Sloman, Hayman, Ohta, Law and Tulving (1988, exp 2) found word fragment completion was primed at above chance levels by study of a list of words 16 months previously. However, while the study list was only presented once, study instructions encouraged a very thorough encoding; participants were asked to copy the study list words onto a piece of paper, rate the familiarity of each word and prepare for a memory test on the words. In a cued recall test of explicit memory Runquist (1983, 1986, 1987) showed that participants could recall words up to two weeks after an initial brief exposure to words when associative cues are given. However, in these experiments some study words were presented in multiple testing sessions to assess the effects of retesting on recall of untested items. This may have aided recall of the items tested only once, leading to inflated estimates of long term retention.

Some studies of shorter term retention for briefly studied words have provided evidence of stable memory after an initial period of forgetting (i.e., an asymptote in the retention function). McBride and Doshier (1997) tested implicit and explicit memory with a stem completion task using study-test lags from one minute to one hour. They found that both types of memory declined over the first 15 minutes before stabilizing above chance with implicit performance below explicit performance. However, even though their analysis assumed an asymptote, they doubted their results supported memory permanence, saying "further decline would be visible in hours or days" (p.380). This

concern is reinforced by their acknowledgement of a failure to equate the average position of the different lag conditions. As a result lag conditions occurring later in the experiment might have produced poorer performance due to greater effects from fatigue or proactive interference.

However, this concern does not apply to Rubin Hinton and Wenzel's (1999), which is often cited as the best existing study of the forgetting function, as it tested 300 participants at a larger number of lags than previous studies. They reported that the best quantitative account of their data was given by a function assuming an above chance asymptote, although the longest lag tested was only several minutes. Nevertheless, Rubin et al. shared McBride and Doshier's cautious interpretation, saying "we believe that the asymptote ... represents a decline too small to detect in our experiments or even in experiments with considerably longer delays" (p.1173). Wixted (2004a) summarizes the pervasiveness of this cautious stance, saying that an asymptote at chance performance "seems to be the view of almost everyone who has ever investigated the mathematical form of forgetting" (p.871).

## Experiment

Our experiment attempts to overcome the limitations of previous studies with the aim of determining whether there is permanent retention or gradual but eventually complete loss of memories for briefly studied stimuli. We tested retention over four separate experimental sessions with study-test lags of 1 minute to 1 hour (session one), 1 day, 1 week and 1 month (sessions two to four). Following McBride and Doshier (1997) both explicit memory (stem-cued recall) and implicit priming (stem completion) were tested in a between-subjects design. These tasks have the advantage of equating tests of the two types of memory in every way except the intention to remember.

Context reinstatement (i.e., restoration of the study context during later testing) has been shown to greatly aid explicit recall (Smith & Vella, 2001). Clearly some aspects of context, both external (e.g., subjects attire) and internal (e.g., subject's mood), will change across sessions. Indeed, one of the reasons that performance may decline across sessions yet not decline within a session at longer lags is that the within session asymptote represents "a constant residual of recall until the experimental context changes" (Rubin et al., 1999, p.1173). If this is the case, even if some memories from the first session are permanent measured retention will decline because they are not retrievable due a lack of appropriate context cues during later test sessions.

Although it is practically impossible to reinstate context perfectly we took two measures to promote reinstatement in later sessions. First, prior to the first session, participants' video taped their walk (first person view) from the foyer of the building where the experiment was conducted into the experimental room. The experimenter also made a quick video of the participant sitting at the chair preparing to start the experiment in order to capture aspects such as attire. Second, participants answered questions about the weather,

their surrounds and mood as well as daily activities just prior to commencement of the first session. The answers to these questions were reviewed, and the video tape replayed, prior to later testing sessions.

## Method

Participants ( $n=32$ ) completed four experimental sessions spaced over 28 days under either explicit or implicit completion instructions. All participants were told that the object of the study was to test lexical skill development over time, specifically their ability to rate the frequency with which words occur and their ability to complete stems. The initial experimental session lasted around 2.1 hours and was divided into two sections separated by an 8.6 minute break. Retention in session one was measured at seven roughly exponentially spaced lags (see Table 1). Sessions two, three and four lasted 35 minutes; session two was held 24 hours later, session three was seven days later and the last session occurred 28 days later. The procedure was identical for participants in both implicit and explicit groups except for the stem completion instructions.

Experimental materials were 905 critical and 505 filler stems and 1170 study words. Critical stems had four or more possible completions, one of which corresponded to a study word. Study words corresponding to the critical stems were chosen because they had the lowest completion rate in a pilot study. The pilot study entailed 20 participants completing all 905 critical stems with the first 4-6 letter word that came to mind and without any prior word study. If two words had the same completion rate the word with the lowest frequency was used. The mean study word completion rate in the pilot, which estimated the chance completion rate in the main experiment, was 11.6%.

After the video tape and questionnaire were completed experimental testing, consisting of thirty 4.3 minute study-test cycles, commenced. These cycles comprised 17 trials in which participants chose which of two words occurred more frequently in their linguistic experience. Each pair appeared on the screen for four seconds. The pair ratings task was used to insure consistency of encoding both within and between groups. Next participants performed 26 stem completion trials. On each trial three letters and three trailing underscores were presented and the participants were asked to complete the stem using 1-3 letters to form either a) a word previously seen (explicit,  $n=16$ ) or b) the first word that came to mind (implicit,  $n=16$ ). The stems were on screen for seven seconds.

Table 1: Study-test lag (minutes), number of items tested at each lag and the mean position within the experiment (minutes from commencement) of the mid-points of lags in each lag condition in Session 1.

Lag	1.2	2.4	5.5	9.5	17.3	33.1	64.2
Tests	86	84	84	85	85	84	84
Position	69.8	70.4	70.2	70.4	70.2	70.1	70.1

Table one shows the number of study words tested per lag as well as the mean position for each lag condition in

session one. Critical stems were tested only once and their test positions were interspersed amongst tests of filler items in order to equate the average position of different lag conditions within the experiment. The two shortest lags were within-cycle, whereas all other lags were between cycles. Tests of between-cycle lags were arranged to minimize testing of items from nearby study positions together in order to minimize the degree to which recall of one item could assist recall following tests of items from nearby study positions (Howard & Kahana, 2002). Critical study words and test stems were randomly allocated to within-subject conditions for each participant.

Later sessions began with a review of the video tape and the experimenter reading back both the questions from the context reinstatement questionnaire and the answers given by the participant. All later sessions consisted of five study-test cycles similar to the first session, except that only 13 pairs, consisting of only filler words which were never tested, were used in the frequency rating trials. The first study test cycle in each session was used to reacquaint participants with the task and used only filler items. The stem completion trials in the remaining cycles used stems corresponding to words studied but not tested in the first session. Hence there were 104 critical-stem tests in each of the later sessions. The study positions of these critical stems were distributed evenly across the first session.

## Results

Bayesian Hierarchical analysis (Shiffrin, Lee Wagenmakers & Kim, 2008) was used to obtain estimates of the population probability of implicit and explicit completion with the studied stem. This analysis assumed a binomial distribution for study completion frequencies and treated subjects as a random effect. The logit of the binomial probability was assumed to have a multivariate normal distribution over subjects that allowed for correlations among the 10 study-test lags (seven in the first session and one for each of the three additional sessions).

The multivariate normal population distribution means were given vague normal priors (precision = 0.001) and its variance-covariance matrix was given a Wishart prior (main diagonal=1, off diagonal=0; Rouder, Lu, Sun, Speckman, Morey & Naveh-Benjamin, 2007). WinBUGS (Lunn, Thomas, Best & Spiegelhalter, 2000) was used to obtain a single chain of 300,000 independent samples from the posterior using a burn in period of 50,000 iterations and thinning that accepted a sample every 50th iteration. Visual inspection of the samples confirmed convergence and independence was confirmed by an examination of the autocorrelation function of the samples. Note that similar results were obtained from the Bayesian analysis and from standard maximum likelihood analysis of each participant. The advantage of the Bayesian approach is that it enabled hierarchical estimation of population distributions. Population retention distributions at the longest lag are of particular interest as they address the issue of whether performance remains generally above chance.

Figure one shows posterior mean estimates and 95% credible intervals on the study-completion-probability scale for the seven lags in the first session. The 95% credible intervals were estimated by the range between the 2.5% and 97.5% quantiles of the posterior mean samples. Explicit performance was generally higher than implicit performance and both declined monotonically for the first five lags before becoming relatively stable at a level well above the chance completion rate. The advantage for explicit performance was much larger up to the third lag and then was relatively constant for the remaining lags.

Figure two shows implicit and explicit performance for the last lag of the first session and the three remaining sessions. Performance was lower in session two relative to the end of session one, with the drop being larger for the explicit condition (0.071) than the implicit condition (0.031). As a result there was little difference in explicit and implicit performance at one day. Both implicit and explicit performance dropped by a similar amount (0.033) from one day to one week. There was little change in performance from one week to one month, with explicit performance rising very slightly (by 0.004) and implicit performance declining very slightly (by 0.005).

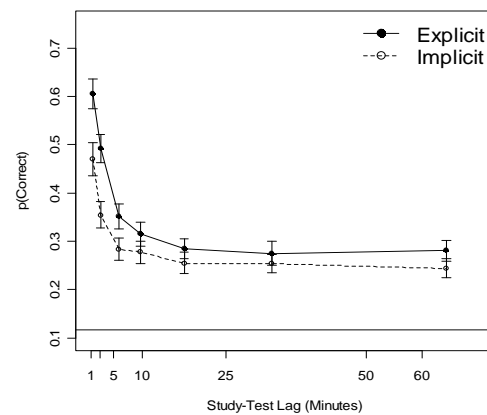


Figure 1: Hierarchical Bayesian estimates of the population proportion of study completions for each of the first session lags with 95% credible intervals. The solid horizontal line is the chance completion level.

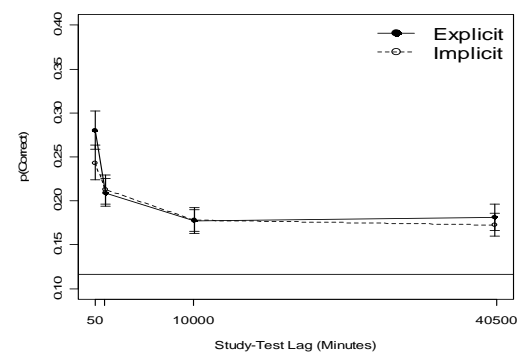


Figure 2: Hierarchical Bayesian estimates of the population proportion of study completions for lags 7 (end of session one) to 10 (28 days post study) with 95% credible intervals. The solid horizontal line is the chance completion rate.

Figure three shows the mean difference in posterior samples between each lag and the lag immediately following it for both the explicit and implicit conditions. The error bars represent a 95% credible interval on the difference distribution. A reliable difference in performance between two adjacent lags is indicated by error bars not crossing the horizontal line marking a zero difference. There are clear differences in performance between the first and second and the second and third lag in both conditions (i.e., the first and second difference points). Beyond that differences are small and for the last two differences in session one are almost exactly zero on average. There is a clear drop in performance in the explicit condition between the end of session one and session two which is not evident in the implicit condition. The decrease from one day to one week (eighth difference point) is close to reliable in both the explicit (2.5<sup>th</sup> percentile = -.0057) and implicit (2.5<sup>th</sup> percentile = -.0028) condition but from one week to 28 days performance is clearly stable for both implicit and explicit tasks.

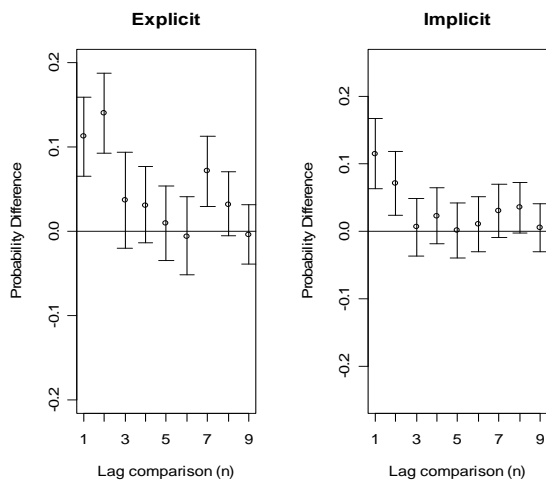


Figure 3: Mean difference between posterior mean samples for lag (n+1)-lag (n). Bars represent 95% credible intervals.

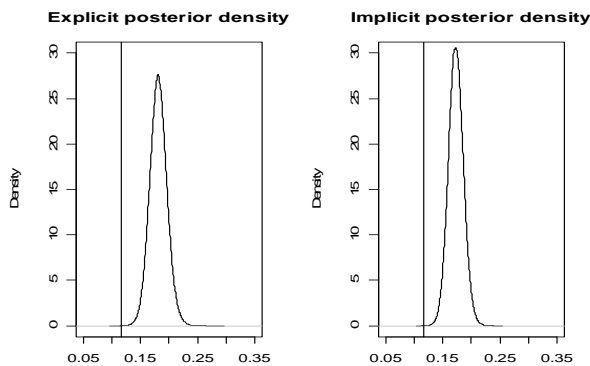


Figure 4: Posterior distributions for mean population study completion probability at 28 days for explicit (left) and implicit (right) tasks. The vertical line represents the chance completion rate.

Figure four shows the posterior distribution for performance at 28 days in both the explicit and implicit conditions. The chance completion rate falls at the  $3 \times 10^{-5}$  quantile of the explicit distribution and the  $2.5 \times 10^{-5}$  quantile of the implicit distribution, indicating that performance was well above chance levels at 28 days. Difference sample t-tests were conducted on the binomial probability of completion at 28 days for both the explicit and implicit conditions. Differences were found in both conditions, explicit  $t(15) = 5.69, p < .001$ ; implicit  $t(15) = 10.54, p < .001$  confirming the results from the posterior distribution (figure 4).

#### Discussion

The results of the experiment clearly show that memory in both stem completion and stem-cued recall tasks is stable, at above chance rates, from one week to 28 days. Performance at 28 days for both tasks was well above chance levels. Although performance on the explicit task was greater than on the implicit task at all lags in the first session, replicating McBride and Doshier's (1997) findings, there were no differences in the latter sessions. Given that the study trial instructions and timing of study presentations (2 seconds per word) allowed only a cursory encoding, the clearly above chance performance at 28 days is a remarkable result. The almost identical performance between explicit and implicit conditions at later lags implies that conscious attempts at recall are not necessary for successful recall, although conscious recall does benefit performance at shorter lags. It also suggests that rehearsal during the study-test lag is unlikely to explain retention performance, as might have been the case in previous studies of long-term retention.

It is possible that the implicit condition may have been contaminated by conscious recall. We examined this possibility by running a further "speeded-implicit" condition which was the same as the implicit condition except that participants were asked to respond as quickly as possible during stem completion. If the first key press took longer than 1.5 seconds a visual "TOO SLOW" warning was given (c.f. Ratcliff & Rouder, 1998) and participants were generally compliant in responding quickly. It has been previously argued that speeded responses limit the use of conscious processes and are therefore a measure of implicit processes (Wilson & Horton, 2002). Although full results for the speeded-implicit condition cannot be reported here due to lack of space, they were very similar to the implicit condition, suggesting that the results reported here are fairly process-pure.

Wixted (2004*ab*) concluded that memories eventually degrade completely. This degradation is driven largely by a non-specific retroactive interference process that disrupts memory consolidation in continuous fashion as storage time increases. The negligible decreases in performance from about 15 minutes to 1 hour and between seven days and 28 days are not consistent with a continuous degradation process. This conclusion is graphically illustrated in Figure 5, which plots the data from all sessions on one graph using logarithmic scale on the abscissa.

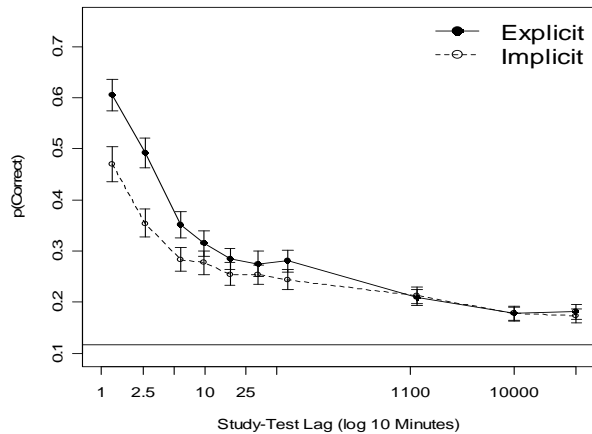


Figure 5: Results for all sessions.

Our data also suggests that there was no further forgetting after seven days. This finding, and the above chance retention at 28 days in both conditions, support Chechile's (2006) suggestion that forgetting functions should permit an asymptote parameter that allows for some memories to be treated as permanent. A strong interpretation of these findings is that they support at least some memories being permanently stored. However, permanence is almost impossible to prove in practice, as the possibility remains open that memory performance could fade when measured over a period longer than the 28 day maximum used here or indeed even any longer period used in another study. One potential approach to this issue, which we are pursuing in on-going work, is to model the data with quantitative forgetting functions and examine the degree of deviation from functions that assume a chance asymptote.

A potential theoretical explanation for above chance asymptotes in long term retention is that some if not all memories, once encoded, are permanent. In this view much of forgetting occurs because memories cannot be retrieved due to lack of effective cues at test. Such retrieval cues have been suggested to mainly consist of aspects of the study context that were encoded with the to-be-retrieved memory (Tulving & Thomson, 1973).

In our experiment the stem is obviously one very powerful cue that co-occurred with the to-be-retrieved memory, perhaps explaining in part the remarkable retention which we observed. Some further aspects of the test environment were also identical between study and test, such as the experimental room, but other aspects, such as participants' mood, had likely changed. Replaying the video tape made at the start of the first session and rehearsing the questions and answers in the questionnaire might have further facilitated retrieval by reinstating these cues, although as we did not manipulate this factor the degree to which it was effective in our experiment remains in question.

While not tested the reinstatement of study session environmental context as a means of improving explicit recall is a robust phenomenon reported in many

experimental paradigms (e.g., Cousins & Hanley, 1996; Gordon & Baddely 1975, 1980, Murnane, Phelps, & Malmberg, 1999). Smith and Vella's (2001) meta-analysis of the context dependent memory literature strongly suggested that both physical and mental reinstatement of study context improves explicit performance. Moreover, and of particular importance to the current study, the analysis indicated that reintroducing context after longer delay periods (>24 hours) leads to a larger effects than at short delay periods ( $\leq 24$  hours). Taken together, this evidence suggests that context reinstatement will aid explicit performance at longer delays.

Implicit memory, however, is defined as increased performance without reference to any learning episode and, therefore, is presumably less reliant on contextual cues. Given this, the context reinstatement account might be thought not to explain retention performance in the implicit condition. However, researchers examining the relationship between putative implicit and explicit memory systems have suggested that they may share some processes and influences. For example, Jacoby and Dallas (1981) proposed that perceptual priming and recognition rely, to differing degrees, on perceptual fluency. Zeelenberg, Pecher, Shiffrin, and Raaijmakers (2003) showed that priming is influenced by cue and target similarity in a manner not dissimilar to explicit memory. Sloman et al. (1988) suggested of their long term priming data that "Reinstatement of the study context at the time of the test...should facilitate performance" (p.238).

Hence, it is possible that implicit performance can be aided by reinstatement of study session context at least to some degree. Alternatively the similarities in our results for implicit and explicit tasks might be taken as evidence for the single system view, where performance differences between explicit and implicit memory tasks are attributed to the tasks rather than the underlying memory system (e.g., Kinder & Shanks, 2001).

In summary, the data presented here suggest that with sufficient retrieval cues memory performance is stable and potentially very long lasting, even if encoding is cursory. Our results are based on a design that has overcome the limitations of previous studies. In short, the design allowed testing of more critical stems per lag, over longer time frames, and with more control over interference levels between lags than has previously been studied. We are currently extending the analyses presented here to examine the implications of our data for quantitative models of the retention function. Rubin et al. (1999) found evidence not only for an asymptote but also for two time scales of forgetting within a session, which they identified with short and long term memory. The results presented in Figure 5 seem to indicate that a third time scale and second asymptote might be needed to accommodate between-session forgetting.

## Acknowledgments

This research was supported by funding from the John and Daphne Research Endowment Fund.

## References

- Bahrick, H. P. (1984). Semantic memory content in permastore: Fifty years of memory for Spanish learned in school. *Journal of Experimental Psychology: General*, *113*, 1-26.
- Chechile, R.A.(2006) Memory Hazard Functions: A vehicle for theory development and test. *Psychological Review*, *113*,31-56.
- Cousins, R. & Hanley, H. (1996). The Effect of Environmental Context on Recall and Category Clustering Scores Following Relational and Individual Item Processing: A Test of the Outshining Hypothesis. *Memory*, *4*, 79-90.
- Gorden, D. R., & Baddeley, A. D. (1980). When does context influence recognition memory? *British Journal of Experimental Psychology*, *71*, 99-104.
- Gorden, D. R. & Baddeley, A. D. (1975). Context dependent memory in two natural environments: Land and underwater. *British Journal of Psychology*, *66*, 325-331.
- Howard, M.W. & Kahana, M. J. (2002a). A distributed representation of temporal context. *Journal of Mathematical Psychology*, *46*, 269-299.
- Jacoby, L.L. & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, *110*, 306-340.
- Jacoby, L. L., Toth, J. P., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General*, *122*, 139-154.
- Kinder, A., & Shanks, D.R. (2001). Amnesia and the declarative/non-declarative distinction: A recurrent network model of classification recognition and repetition priming. *Journal of Cognitive Neuroscience*, *13*, 648-669.
- Lunn, D., Thomas, A., Best, N., & Spiegelhalter, D. (2000). WinBUGS a Bayesian modeling framework: concepts, structure, and extensibility. *Statistics and Computing*, *10*, 325-337.
- McBride, D. M., & Doshier, B. A. (1997) A comparison of forgetting in an implicit and explicit memory task. *Journal of Experimental Psychology: General*, *126*, 371-392.
- Murnane, K., Phelps, M. P., & Malmberg, K. (1999). Context dependent recognition memory. *Journal of Experimental Psychology: General*, *128*, 403-415.
- Noice, T., & Noice, H. (2002). Very long term recognition and recall of well learned material. *Applied cognitive psychology*, *16*, 259-272.
- Ratcliff, R. & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*, 347-356.
- Rouder, J.N., Lu, J., Sun,D., Specman,P.L., Morey, R.D., & Naveh-Benjamin (2007). Signal detection models with random participants and item effects. *Psychometrika*, *72*, 621-642.
- Rubin, D.C., Hinton, S., & Wenzel, A.E. (1999). The precise time course of forgetting *Journal of Experimental Psychology*, *25*, 1161-1176.
- Runquist, W. (1983). Some effects of remembering on forgetting. *Memory and Cognition*, *11*, 641-650.
- Runquist, W. (1986). The effect of testing on the forgetting of related and unrelated associates. *Canadian Journal of Psychology*, *40*, 65-76.
- Runquist, W. (1987) Retrieval Specificity and the attenuation of forgetting by testing. *Canadian Journal of Psychology*, *41*, 84-90.
- Schmidt, H. G., Peck, V. H., Paas, F., & van Breukelen, G. J. P. (2000). Remembering the street names of ones childhood neighborhood: A study of very long term retention. *Memory*, *8*, 37-49.
- Sloman, S.A., Hayman,C.A.G. ,Ohta, N., Law, J., & Tulving, E. (1988). Forgetting in Primed Fragment Completion. *Journal of Experimental Psychology; Learning Memory and Cognition*, *14*, 223-239.
- Shiffrin, R. M., Lee, M. D., Kim, W., & Wagenmakers, E.-J. (2008). A survey of model evaluation approaches with a tutorial on hierarchical Bayesian methods. *Cognitive Science*, *32*, 1248-1284.
- Smith, S. M., & Vella, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin & Review*, *8*, 203-220.
- Squire, L. R. (1994). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. In D. L. Schacter & E. Tulving (Eds.), *Memory Systems 1994* (pp. 203-231). Cambridge, MA: MIT Press.
- Squire, L. R. (1989). On the course of forgetting in very long term retention. *Journal of Experimental Psychology: Learning Memory and Cognition*, *15*, 241-245.
- Tulving, E., & Thomson, D.M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*, 352-373.
- Wilson, D. E. & Horton, K. D. (2002). Comparing techniques for estimating automatic retrieval: Effects of retention interval. *Psychonomic Bulletin & Review*, *9*, 566-574.
- Wixted, J. T. (2004 a). On common ground: Jost's (1897) law of forgetting and Ribots's (1881) law of retrograde amnesia. *Psychological Review*, *111*, 864-879.
- Wixted, J. T. (2004 b) The psychology and neuroscience of forgetting. *Annual Review of Psychology*, *55*, 235-269.
- Zeelenberg, R., Pecher, D., Shiffrin, R.M., & Raaijmakers, J.G.W. (2003). Semantic context effects and priming in word association. *Psychonomic Bulletin and Review*, *10*, 653-660.