

Self-evaluation as a moderating factor of strategy change in directed forgetting benefits

LILI SAHAKYAN

Florida State University, Tallahassee, Florida

PETER F. DELANEY

University of Florida, Gainesville, Florida

and

COLLEEN M. KELLEY

Florida State University, Tallahassee, Florida

In list method directed forgetting, instructing people to forget a studied word list usually results in better recall for a newly studied list. Sahakyan and Delaney (2003) have suggested that these benefits are due to a change in encoding strategy that occurs between the study of the first list and the study of the second list. To investigate what might mediate such strategy change decisions, in two experiments we induced both *forget* and *remember* participants to evaluate their memory performance on the two lists. In Experiment 1, they were asked to explicitly recall the items from the first list before studying the second list. In Experiment 2, after the study of the first list, the participants provided a rapid aggregate judgment of learning. Evaluation eliminated the differences between the forget and remember groups for the second list in both experiments, because the remember group achieved recall levels comparable to those for the forget group. The role of performance evaluation in mediating directed forgetting benefits is discussed.

In the typical list method directed forgetting experiment, participants receive two lists of words to study for a later memory test. After studying the first list, some participants are instructed to forget the prior items, whereas other participants are told to keep remembering the prior items. Both groups then receive the second list, which is followed by an instruction to remember all studied items. Individuals instructed to forget the first list recall the second list better than people who are not given the forget instruction—a finding referred to as the *benefits* of directed forgetting (E. L. Bjork & R. A. Bjork, 1996; R. A. Bjork, 1970; Geiselman, R. A. Bjork, and Fishman, 1983; Liu, 2001; Reitman, Malin, R. A. Bjork, & Hingman, 1973). In addition, people's ability to recall the to-be-forgotten items on the first list is impaired in comparison with the recall of those who are not given the forget instruction—a finding known as the *costs* of directed forgetting (Geiselman et al., 1983; Liu, 2001; Reitman et al., 1973).

One interpretation of directed forgetting is based on retrieval inhibition (e.g., E. L. Bjork & R. A. Bjork, 1996; R. A. Bjork, 1989; Geiselman et al., 1983). According to that account, the instruction to forget initiates a process that blocks or inhibits access to List 1 items during recall,

resulting in poorer retrieval (hence, the costs). The benefits of directed forgetting have been attributed to escape from proactive interference due to the reduced accessibility of the first list (E. L. Bjork & R. A. Bjork, 1996; R. A. Bjork, 1989; R. A. Bjork & Woodward, 1973). Recall of the second list by the forget group is comparable to recall in a condition that receives only a single list to study.

Sahakyan and Kelley (2002) have proposed an alternative explanation for the costs and benefits of directed forgetting, based on participants' retrospective reports. Their participants reported that, after hearing the forget instruction, they attempted to think of something completely different from the first list, which created a new mental context. The second list was thus subsequently encoded in the new mental context, and the final recall of both lists occurred in the new mental context. Thus, at recall the mental context mismatches the context present during encoding of the first list, producing the costs of directed forgetting. The benefits of directed forgetting follow from reduced proactive interference due to encoding the two lists in different contexts. In accord with the context change hypothesis, Sahakyan and Kelley demonstrated effects similar to those of directed forgetting when they simply induced a mental context change in a group instructed to remember both lists of words by requiring participants to think of some unrelated activity between the two study lists. This *remember plus context change* condition showed costs and benefits comparable

Correspondence concerning this article should be addressed to L. Sahakyan, Department of Psychology, University of Florida, Box 112250, Gainesville, FL 32611-2250 (e-mail: sahakyan@ufl.edu).

to those regularly observed in the *forget* condition. In addition, mentally reinstating the original learning context before recall reduced the magnitude of directed forgetting costs and benefits predicted by the contextual change account.

Liu (2001) found that a delay before recall in directed forgetting reduced the costs of directed forgetting but did not affect the benefits. This suggests that the benefits of directed forgetting stem from better encoding of the second list. In accord with this hypothesis, Sahakyan and Delaney (2003) analyzed retrospective reports of study strategies from participants in the Sahakyan and Kelley (2002) directed forgetting condition. The participants in the forget and remember plus context change conditions reported switching to a superior encoding strategy on the second list more often than did the participants in the regular remember condition. The benefits appeared to be due to participants' changing to deep encoding strategies more frequently in the forget condition and in the remember plus context change condition than in the remember condition. When participants' encoding strategies were kept consistent between the two lists by experimental instructions, the benefits of directed forgetting disappeared, although the costs remained (Sahakyan & Delaney, 2003).

An unanswered question is *why* there would be a higher rate of strategy change in the forget and remember plus context change groups than in the remember group. The memory strategy choice literature contains cases in which people continue to rely on the same inefficient strategy throughout an experiment even when free to choose a strategy of their own, and also cases in which people switch to more effective encoding strategies across trials (Jacoby, 1973; Pressley, Levin, & Ghatala, 1984; Shaughnessy, 1981). Participants' failure to switch to a better encoding strategy during study is usually attributed to participants' inability to monitor the efficacy of strategies during the execution of the strategy and to participants' lack of knowledge about the relative efficacy of encoding strategies (Pressley et al., 1984; Shaughnessy, 1981). People's choice between two study strategies is essentially random unless they have an opportunity to practice using the strategies and receive feedback about their effectiveness (Biemann-Copland & Charness, 1994; Dunlosky & Hertzog, 2000; Pressley et al., 1984). Interestingly, instructions to evaluate one's own expected performance also led participants to choose more effective strategies (Pressley et al., 1984). In directed forgetting, the experimenters give no feedback about the effectiveness of encoding strategies, so some form of self-generated feedback following the study of List 1 most likely motivates any change of strategy between study of the first list and study of the second list.

Our hypothesis is that when participants in the remember group are engaging in their typical maintenance rehearsal strategy to encode List 1, they rarely stop to reflect on its efficacy. The cost of interrupting rehearsal to evaluate strategy effectiveness may seem too high to participants. However, when the first list is followed by in-

structions to forget that list and prepare to study the "real" list, participants are likely to evaluate their List 1 performance in preparation for List 2. Doing so leads many to realize that their strategy is not very fruitful, and this motivates a change to a better strategy on List 2. The remember plus context change instructions in Sahakyan and Kelley (2002) may provide a similar segmentation of List 1 and List 2 into two separate events and so provoke evaluation of List 1 performance and a shift to a better strategy on List 2. In contrast, in the condition where List 1 is treated as the first half of the to-be-studied materials, participants are less likely to evaluate how they are doing, and they simply continue with the same encoding strategy for List 2.

EXPERIMENT 1

Experiment 1 tested the standard directed forgetting forget and remember instruction groups with or without evaluation instructions. Participants in the evaluation condition were asked to recall List 1 immediately after study but before receiving the forget or remember cue. We predicted that if strategy changes brought on by self-evaluation were the usual cause of benefits, any manipulation that provided participants with direct feedback about the effectiveness of their encoding strategies ought to improve List 2 performance. We therefore expected that the extra retrieval opportunity would eliminate the benefits, because all participants in the evaluation condition irrespective of the forget or remember cue would be likely to improve their performance on List 2. Thus, we predicted benefits in the standard forget versus remember comparison, but no benefits in the forget plus evaluation versus remember versus evaluation comparison.

Method

Participants. The participants were 160 Florida State University undergraduates who participated for course credit. They were tested in groups of 4 to 6, with 40 participants in each of the four experimental conditions.

Materials and Design. Two lists of 16 unrelated, medium-frequency English nouns were drawn from the Kučera and Francis (1967) norms. Order of presentation of lists was counterbalanced, as was order of list recall at final test. The design of the experiment was a 2 (cue: remember or forget) \times 2 (group: evaluation or control) between-subjects design.

Procedure. Words were presented on each participant's computer screen at a rate of 5 sec/word. After studying List 1, the participants in the evaluation condition had 60 sec to recall the words in any order. Control participants spent those 60 sec solving three-digit arithmetic problems on paper. Immediately afterward, the participants in both the evaluation and the control conditions were verbally presented with either the forget or the remember instructions. The forget instructions specified that the first list was only for practice to familiarize the participants with the task and that there was no need to remember those items. The remember instructions specified that the first list included only the first half of the items from the study list and that those items needed to be remembered for a later memory test. Then all participants studied List 2, which was again followed by 60 sec to recall List 2 or to do arithmetic. List 2 was always followed by the instruction to remember it. All participants then received a

90-sec filler task before the final recall test. At test, half of the participants in each condition were told to recall the first list on one sheet and then the second list on another sheet (with 60 sec/list to recall), while the other half recalled the lists in the opposite order.

Results and Discussion

In all analyses, the dependent measure was the proportion of words correctly recalled from the appropriate word list. Initial recall performance in the evaluation condition was .49 for both forget and remember groups on List 1, and .51 and .48 respectively for List 2.

The directed forgetting costs and benefits were analyzed separately (see Figure 1). For costs, a two-way analysis of variance (ANOVA) with group (control vs. evaluation) and cue (forget vs. remember) was computed on the proportion of final List 1 recall. There was a main effect of cue [$F(1,156) = 17.01, MS_e = .024, p < .001$], with the remember participants recalling significantly more words than the forget participants (.28 vs. .18). The

effect of group was marginally significant [$F(1,156) = 3.74, MS_e = .024, p = .06$], with the evaluation condition recalling more words than the control condition (.25 vs. .21). This is consistent with the large body of prior research on the effects of retrieval practice on future recallability of the items. The interaction was not significant ($F < 1$). To summarize, the expected costs emerged in both the control and the evaluation conditions.

For the analysis of directed forgetting benefits, the same cue \times group ANOVA was computed on the proportion of final List 2 recall. There was a significant interaction of cue and group [$F(1,156) = 3.97, MS_e = .021, p < .05$]. Simple effects analyses revealed that in the control condition, the forget group recalled significantly more of List 2 than did the remember group, showing the standard directed forgetting benefits [$F(1,156) = 8.74, MS_e = .021, p < .01$]. However, there was no difference between the forget and remember groups in the evaluation condition ($F < 1$). Thus, as predicted, significant directed

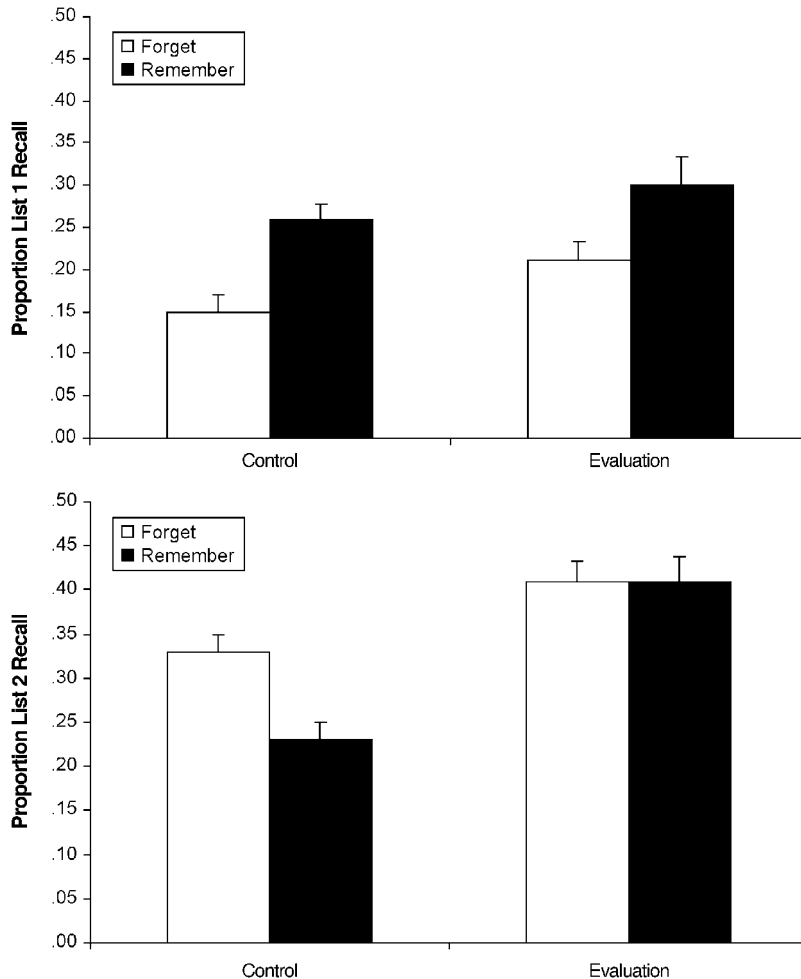


Figure 1. Proportion of words recalled from List 1 (top panel) and List 2 (bottom panel) by experimental condition, Experiment 1. Error bars represent $\pm SE$.

forgetting benefits were found only in the control group, and not in the evaluation condition. We interpret this as reflecting that retrieval of List 1 immediately after studying it led the remember group to improve its List 2 performance relative to List 1 performance [$F(1,39) = 20.74, MS_e = .01, p < .001$], which was not the case in the control group ($F < 1$).

EXPERIMENT 2

The results of Experiment 1 suggested that people responded to the performance assessment induced by the initial recall of List 1 in a way that improved memory for the subsequent list, presumably by adopting a more effective study strategy on the second list. However, participants in directed forgetting studies are unlikely to attempt an exhaustive recall of List 1 to evaluate their encoding strategy, because the interval between the two lists is quite short. Therefore, we propose that a more heuristic approach to self-evaluation prompts the strategy change decisions. Monitoring one's knowledge could take the form of a metacognitive assessment through judgments of learning (see, e.g., Koriat, 1997; Mazzoni & Nelson, 1995) or feeling of knowing judgments (e.g., Koriat, 1995; Schwartz & Metcalfe, 1992).

In Experiment 2, aggregate judgments of learning (JOLs) provided the performance evaluation manipulation. In aggregate JOLs, people study the whole list and then predict the total number of items they will recall. Mazzoni and Nelson (1995) showed that aggregate JOLs are rather good predictors of recall in intentional learning tasks. We therefore asked some participants to provide aggregate JOLs after studying the first list but before receiving the instruction to forget or remember.

One reason why we used JOLs was to address further whether evaluation mediates the effects of the forget cue on List 2 recall. If the forget cue typically produces benefits for List 2 recall by some mechanism other than evaluation, the addition of evaluation should increase benefits in comparison with those for the standard forget condition. In Experiment 1, the evaluation manipulation of initial recall of List 1 and List 2 increased recall in comparison with the standard study conditions, so we could not compare List 2 recall for the standard forget versus the forget plus evaluation condition. However, providing aggregate JOLs should not inflate recall rates at final test, so in Experiment 2 we could directly compare performance in the forget plus JOLs and standard forget conditions. We expected that if the usual pattern of benefits was caused by some factor other than forget cue induced evaluation, evaluation would lead to better List 2 performance for *both* the remember and the forget participants relative to their corresponding no-JOL group. Alternatively, if evaluation caused benefits in the directed forgetting paradigm, then the evaluation forget group should not show any additional recall advantage in comparison with the standard forget group, whereas the eval-

uation remember group should show better recall than the standard remember group would.

Method

Participants. The participants were 136 FSU undergraduates who participated for course credit. They were tested individually, with 34 participants in each of the four experimental conditions.

Materials and Design. Two new lists of 15 unrelated English nouns of medium frequency were drawn from Kučera and Francis's (1967) norms. The design of the experiment was cue (forget or remember) \times group (evaluation or control), between subjects.

Procedure. We repeated the procedure of Experiment 1, with the following exceptions. In the *control* group, the delay between presentations of the two lists was reduced to 15 sec, on the basis of pilot studies that determined how long it took to provide aggregate JOLs. The delay was unfilled. During that time, the *evaluation* group gave aggregate JOLs by predicting how many words from List 1 they would be able to recall at the final test. In each case the instruction to remember or forget immediately preceded the presentation of List 2. We controlled output order, with List 1 recall preceding List 2 recall.

Results and Discussion

Proportion correct List 1 recall was analyzed using a two-way ANOVA with the factors cue (forget vs. remember) and group (evaluation vs. control) (Figure 2, top panel). There was a main effect of cue [$F(1,132) = 12.37, MS_e = .036, p < .001$], confirming the costs of directed forgetting. The remember group recalled more words from the first list than did the forget group (.36 vs. .24). Neither the main effect of group nor the interaction was significant (both $F_s < 1$).

The same variables were used to analyze proportion correct List 2 recall (Figure 2, bottom panel). The main effect of cue was significant [$F(1,132) = 4.17, MS_e = .028, p < .05$], with better recall in the forget condition than in the remember condition (.42 vs. .36). There was also a significant main effect of group [$F(1,132) = 4.50, MS_e = .028, p < .05$], showing that the evaluation group recalled more words than did the control group (.43 vs. .36). The interaction term approached but did not reach significance [$F(1,132) = 2.30, MS_e = .028, p = .13$]. To test the primary predictions regarding the benefits, we carried out two planned comparisons. The first contrast was between the remember and forget groups of the control condition, and was significant [$F(1,132) = 6.32, MS_e = .028, p < .01$]. As predicted, the forget group of the control condition recalled significantly more words than did the remember group (.42 vs. .31), showing the standard benefits of directed forgetting. The second contrast tested the same groups in the evaluation condition, and was not significant ($F < 1$).

As a further test of the evaluation interpretation of forget cue benefits, we compared the standard forget condition and the forget plus evaluation condition. They performed similarly on List 2 ($F < 1$), suggesting no additional benefit for the forget plus evaluation condition. However, the participants in the remember plus evaluation condition recalled significantly more words than the

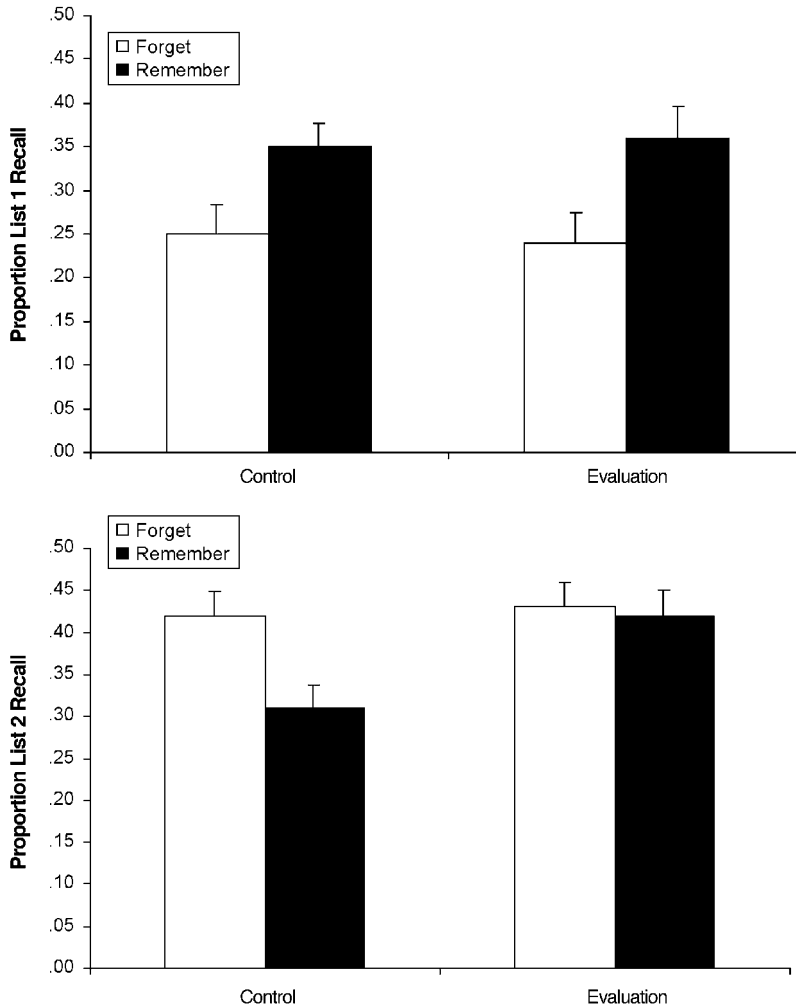


Figure 2. Proportion of words recalled from List 1 (top panel) and List 2 (bottom panel) by experimental condition, Experiment 2. Error bars represent $\pm SE$.

participants in the standard remember condition did [$F(1,66) = 6.68, MS_e = .028, p < .05$], suggesting that the norm for the remember condition is to not engage in evaluation after List 1.

GENERAL DISCUSSION

These experiments provide further tests of the hypothesis that the benefits of being instructed to forget a first list in preparation for studying a second list are mediated by participants' evaluation of first list performance and a shift to better encoding strategies. Evaluation induced by initial retrieval of the lists (Experiment 1) or aggregate JOLs (Experiment 2) reduced the benefits of the directed forgetting instructions by raising performance of participants in the remember condition. If the forget cue operated by some mechanism independent of evaluation, we would expect that participants ought to perform better on List 2 in the forget plus evaluation

condition than in the standard forget condition. However, the forget instruction did not improve recall beyond the effects of evaluation induced by JOLs in Experiment 2. We interpret these results as indicating that the benefits of the forget instruction operate by inducing evaluation of List 1 performance and a shift to a better encoding strategy for List 2.

Converging evidence for the strategy shift hypothesis comes from an analysis of retrospective reports of strategies in a directed forgetting experiment (Sahakyan & Delaney, 2003). Participants in the forget group reported that they switched from a shallow to a deep encoding strategy between lists. Such switching happened sufficiently often to account for the benefits in list method directed forgetting. The strategy shifts were triggered by both contextual change instructions and forget instructions.

Additional converging evidence for the strategy shift account of the benefits of directed forgetting comes from a study in which the experimental instructions controlled

what strategy participants used on both List 1 and List 2. In accord with a strategy shift hypothesis, when encoding strategy was controlled, the benefits of directed forgetting were eliminated (Sahakyan & Delaney, 2003). Controlling encoding strategies did not affect the costs of directed forgetting, however. Although the benefits of directed forgetting may be a result of strategy shifts, the costs may be due to internal shifts of context between lists (Sahakyan & Kelley, 2002).

One might argue a change in encoding strategy between List 1 and List 2 constitutes a context change, and it is the effect of context change that leads to superior List 2 performance rather than the strategy change. A change in encoding strategy could be considered part of context, broadly defined. However, by that reasoning, we would predict that participants in the remember group who are asked to evaluate their List 1 performance (and so change to a better strategy on List 2) would show the costs typical of directed forgetting—that is, to show lower List 1 recall than the standard remember group. Contrary to that hypothesis, we saw equivalent List 1 performance in the remember plus evaluation condition in comparison with the standard remember condition ($F < 1$).

It is unclear whether other theories of the benefits of directed forgetting, such as retrieval inhibition (R. A. Bjork, 1989), selective rehearsal of List 1 across the forget and remember groups (MacLeod, Dodd, Sheard, Wilson, & Bibi, 2003), or the mental context change hypothesis (Sahakyan & Kelley, 2002) can accommodate the results of the present experiments or those of Sahakyan and Delaney (2003). Evaluation of List 1 might lead to better encoding of List 1, which might override the effects of retrieval inhibition, leading to reduced costs and benefits, but that is not what we found. The evaluation of List 1 performance effectively creates rehearsal of List 1 even in the forget condition. If selective rehearsal is responsible for the costs and benefits, evaluation might have reduced costs and benefits in the forget condition by raising List 1 performance for forget participants, but that is not what we found. If strategy change between List 1 and List 2 contributes to a greater change of context between List 1 and List 2, we should have found greater costs in the evaluation conditions, but that is not what we found.

In other theories of directed forgetting, the costs and benefits of directed forgetting stem from the same process. However, we have found dissociations between costs and benefits consistent with the idea that the benefits stem from an encoding strategy shift on List 2, whereas the costs stem from a different mechanism (see also Liu, 2001). The costs of directed forgetting can still be accounted for by context change theory, retrieval inhibition, or differential rehearsal. However, evidence is accumulating that the benefits of directed forgetting may reflect better encoding of items on the second list.

REFERENCES

- BIEMAN-COPLAND, S., & CHARNESS, N. (1994). Memory knowledge and memory monitoring in adulthood. *Psychology & Aging, 9*, 287-302.
- BJORK, E. L., & BJORK, R. A. (1996). Continuing influences of to-be-forgotten information. *Consciousness & Cognition, 5*, 176-196.
- BJORK, R. A. (1970). Positive forgetting: The noninterference of items intentionally forgotten. *Journal of Verbal Learning & Verbal Behavior, 9*, 255-268.
- BJORK, R. A. (1989). Retrieval inhibition as an adaptive mechanism in human memory. In H. L. Roediger III & F. I. M. Craik (Eds.), *Varieties of memory and consciousness: Essays in honour of Endel Tulving* (pp. 309-330). Hillsdale, NJ: Erlbaum.
- BJORK, R. A., & WOODWARD, A. E. (1973). Directed forgetting of individual words in free recall. *Journal of Experimental Psychology, 99*, 22-27.
- DUNLOSKY, J., & HERTZOG, C. (2000). Updating knowledge about encoding strategies: A componential analysis of learning about strategy effectiveness from task experience. *Psychology & Aging, 15*, 462-474.
- GEISELMAN, R. E., BJORK, R. A., & FISHMAN, D. L. (1983). Disrupted retrieval in directed forgetting: A link with posthypnotic amnesia. *Journal of Experimental Psychology: General, 112*, 58-72.
- JACOBY, L. L. (1973). Encoding processes, rehearsal, and recall requirements. *Journal of Verbal Learning & Verbal Behavior, 12*, 302-310.
- KORIAT, A. (1995). Dissociating knowing and the feeling of knowing: Further evidence for the accessibility model. *Journal of Experimental Psychology: General, 124*, 311-333.
- KORIAT, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General, 126*, 349-370.
- KUCERA, H., & FRANCIS, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- LIU, X. (2001). *On the dynamics of directed forgetting: Facilitation and interference in the updating of human memory*. Unpublished doctoral dissertation, University of California, Los Angeles.
- MACLEOD, C. M., DODD, M. D., SHEARD, E. D., WILSON, D. E., & BIBI, U. (2003). In opposition to inhibition. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 43, pp. 163-214). New York: Elsevier, Academic Press.
- MAZZONI, G., & NELSON, T. O. (1995). Judgments of learning are affected by the kind of encoding in ways that cannot be attributed to the level of recall. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21*, 1263-1274.
- PRESSLEY, M., LEVIN, J. R., & GHATALA, E. S. (1984). Memory strategy monitoring in adults and children. *Journal of Verbal Learning & Verbal Behavior, 23*, 270-288.
- REITMAN, W., MALIN, J. T., BJORK, R. A., & HINGMAN, B. (1973). Strategy control and directed forgetting. *Journal of Verbal Learning & Verbal Behavior, 12*, 140-149.
- SAHAKYAN, L., & DELANEY, P. F. (2003). Can encoding differences explain the benefits of directed forgetting in the list method paradigm? *Journal of Memory & Language, 48*, 195-206.
- SAHAKYAN, L., & KELLEY, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 28*, 1064-1072.
- SCHWARTZ, B. L., & METCALFE, J. (1992). Cue familiarity but not target retrievability enhances feeling-of-knowing judgments. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 18*, 1074-1083.
- SHAUGHNESSY, J. J. (1981). Memory monitoring accuracy and modification of rehearsal strategies. *Journal of Verbal Learning & Verbal Behavior, 20*, 216-230.