

# Forgetting is effortful: Evidence from reaction time probes in an item-method directed forgetting task

JONATHAN M. FAWCETT AND TRACY L. TAYLOR  
*Dalhousie University, Halifax, Nova Scotia, Canada*

Reaction time (RT) was measured in response to visual detection probes embedded within an item-method directed forgetting paradigm. In Experiment 1, study words were presented individually followed by an instruction to remember (R) or forget (F). Probes were presented at stimulus onset asynchronies (SOAs) of 1,400, 1,800, or 2,600 msec in relation to the study word or memory instruction. After the study trials, a *yes–no* recognition task measured retention of R and F words. Experiment 2 added a no-word control condition and no-probe catch trials. In both experiments, post-F probe RTs were longer than post-R probe RTs at early SOAs. Confirming that participants attended to the memory instructions, there was a significant directed forgetting effect, with greater recognition of R than of F words. These findings contradict the view that directed forgetting in the item-method paradigm is due to the passive decay of nonrehearsed F items; instead, they are consistent with the view that intentional forgetting in an item-method paradigm occurs via the operation of an active, potentially inhibitory, cognitive process.

A typical person is required to retain and retrieve a sizable quantity of information on a daily basis. When retention or retrieval fails, the consequences are unpleasant, and, as a result, forgetting often has a negative connotation. However, forgetting is not always a failure of memory. For example, when performed intentionally, forgetting may prevent irrelevant or outdated information from interfering with the encoding and/or retrieval of relevant and current information (e.g., Bjork, 1970, 1972, 1989; James, 1890/1950). In such a case, far from being an unwanted failure of memory, forgetting underlies efficient memorial function.

In the laboratory, intentional forgetting is studied using a directed forgetting paradigm (for a review, see MacLeod, 1998). There are many variants of the directed forgetting paradigm (for a review, see Basden & Basden, 1998, and Bjork, 1972), most of which may be categorized as either item method (e.g., MacLeod, 1989; Muther, 1965; Paller, 1990; Tekcan & Aktürk, 2001) or list method (e.g., Conway, Harries, Noyes, Racsma'ny, & Frankish, 2000; Geiselman, Bjork, & Fishman, 1983; McNally, Clancy, Barrett, & Parker, 2004). These paradigms are distinguished by the timing of the memory instructions (i.e., remember, R; forget, F) relative to the study items (Bjork, 1972). In item-method tasks, R and F instructions are presented randomly (but with equal probability) on a trial-by-trial basis, usually following the presentation of each study word. In list-method tasks, an instruction to forget is given following the presentation of a discrete set of words (i.e., a

list), after which participants are presented with a second set to remember (Basden & Basden, 1998; Bjork, 1972). Both paradigms are concerned with the effect of memory instruction on subsequent memory performance.

Although it was believed that item- and list-method paradigms were interchangeable, recent findings suggest that the effect of R and F instructions within these paradigms may be mediated by distinct underlying mechanisms. Basden, Basden, and Gargano (1993; see also Bjork, 1989) observed that item-method tasks consistently produce better memory for R than for F words (i.e., a directed forgetting effect) for both recognition and recall, whereas list-method tasks produce such differences only for recall. This observation is inconsistent with a common mechanism and has resulted in research that compares item-method tasks with list-method tasks in an effort to discern the mechanism(s) through which R and F instructions operate to influence later memory performance (e.g., Basden, Basden, Coe, Decker, & Crutcher, 1994; Conway & Fthenaki, 2003; MacLeod, 1999). The culmination of this research suggests that item-method directed forgetting is mediated by selective rehearsal favoring R items (Basden et al., 1993; see also Basden & Basden, 1998), whereas list-method directed forgetting is mediated by the inhibition of F items at the time of retrieval (e.g., Geiselman et al., 1983) or by a change of mental context between study and test (e.g., Sahakyan & Delaney, 2005; Sahakyan & Kelley, 2002; for another alternative, see also Sheard & MacLeod, 2005).

---

J. M. Fawcett, [jmfawcet@dal.ca](mailto:jmfawcet@dal.ca)

---

The item-method paradigm is of interest for the present study. According to the selective rehearsal account, directed forgetting in this paradigm is due to the selective rehearsal of R words. The claim is that, in a typical paradigm that presents the memory instruction following the disappearance of the study word, participants use maintenance rehearsal (see Craik & Lockhart, 1972; Lockhart & Craik, 1990; Woodward, Bjork, & Jongeward, 1973) to refresh each word in working memory until the presentation of an R or F instruction. This rehearsal maintains only the study word and does not differentially affect subsequent memory performance for R or F words; in the event of simultaneous word-instruction presentation (e.g., Paller, 1990), maintenance rehearsal is unnecessary. Regardless of whether the instruction is presented following or simultaneous with the study word, if participants are instructed to remember the word, they initiate elaborative rehearsal (although see Golding, Roper, & Hauselt, 1996) to commit that word to memory; if the instruction is to forget the word, participants drop the word from their rehearsal set. The standard interpretation implies that dropping an F word from the rehearsal set amounts to the passive decay of its representation (cf. Elmes, 1969).

An alternative hypothesis suggests that intentional forgetting in the item-method paradigm depends on attentional inhibition (cf. Hasher & Zacks, 1988; Zacks & Hasher, 1994; although see MacLeod, Dodd, Sheard, Wilson, & Bibi, 2003). Zacks, Radvansky, and Hasher (1996) suggested that receipt of an F instruction engages attentional mechanisms that actively suppress the representation of F words at study. In this way, attention expunges irrelevant items from working memory (e.g., Taylor, 2005; Taylor & Fawcett, 2008) and prevents them from regaining easy access (however, see Marks & Dulaney, 2001). Although such suppression may briefly draw on limited-capacity cognitive resources, these resources are ultimately freed for processing relevant (R) items rather than for irrelevant (F) items (see Roediger & Crowder, 1972).

Zacks et al. (1996) argued in support of the attentional inhibition hypothesis on the basis of a comparison of directed forgetting in younger and older adults. In an item-method task, Zacks et al. (1996, Experiments 1A and 1B) observed that older adults exhibited a smaller directed forgetting effect than did younger adults. This was due to apparent difficulty in suppressing F words: Older adults remembered approximately the same number of F words as did the younger adults, but fewer R words (see also Sego, Golding, & Gottlob, 2006, Experiments 1A and 1B). Although these findings are consistent with an attentional inhibition account, they may also be attributed to age-related difficulties with source monitoring (e.g., Henkel, Johnson, & De Leonardis, 1998) that result in poor selective rehearsal (e.g., older adults cannot recall whether a word was F instructed, and therefore rehearse it instead of allowing its representation to decay). For this reason, Zacks et al.'s results do not constitute strong support for the attentional inhibition hypothesis. The purpose of the present study was to provide a stronger test of this hypothesis.

According to the attentional inhibition hypothesis, a directed forgetting effect arises from the initial active sup-

pression of F words, followed by the focused processing of R words. Thus, forgetting and remembering are both effortful processes. This view contrasts with the passive decay interpretation of the selective rehearsal account, wherein remembering is effortful and forgetting is not (i.e., because forgetting involves passive decay of the now-irrelevant F item representation). The purpose of the present experiments was to test these characterizations by providing evidence regarding the active-passive nature of instantiating an F instruction, as measured at the time of encoding (as opposed to at the time of retrieval). To do this, we integrated a probe-detection task into a standard item-method directed forgetting paradigm. The rationale is that a cognitively demanding task will require greater resource allocation than will a less demanding task. As a result, participants would be slower to detect the probe when it was presented in the context of a more demanding task than when it was presented in the context of one less demanding (e.g., Kahneman, 1973). This reasoning would suggest that it is possible to quantify the relative cognitive demands of R and F instructions. Furthermore, by varying the interval between the memory instruction and the probe, it is possible to observe the time course of these relative cognitive demands.

According to the passive view of intentional forgetting, it should be initially more demanding cognitively to instantiate an R instruction than to instantiate an F instruction. This is because an R instruction engages elaborative rehearsal, whereas an F instruction results in passive decay of the previously maintained word. As a result, initially, probe RTs should be slower following R instructions than following F instructions. It is possible that this RT difference will be maintained at longer stimulus onset asynchronies (SOAs), reflecting consistently greater cognitive load following R than following F instructions. However, it is also possible that at longer postinstruction intervals, cumulative rehearsal of R words from previous trials will lead to equivalent probe RTs on R and F trials (i.e., because cumulative rehearsal would occur on *both* types of trial).

Conversely, according to the active view of intentional forgetting, instantiating an F instruction should be initially at least as cognitively demanding as instantiating an R instruction. This is because both the F and R instructions engage active cognitive mechanisms, one to expunge words from working memory (in the case of an F instruction) and one to elaboratively rehearse words in working memory (in the case of an R instruction). As long as the active mechanism engaged by an R instruction is not *more* demanding than the active mechanism engaged by an F instruction, probe RTs should never be longer following R than following F instructions; if they are, the predictions of the active view become indistinguishable from the predictions of the passive view. Allowing for this caveat, if the active view is correct, probe RTs initially should be equivalent following R and F instructions, and could even be slower following F than following R instructions. At later intervals, it is again possible that cumulative rehearsal of R words from preceding trials will lead to equivalent probe RTs on R and F trials.

## Predictions

Table 1 summarizes predictions for the postinstruction probe RTs for relatively short and relatively long SOAs. The long instruction–probe SOAs are relevant to determining whether there may be cumulative rehearsal of preceding R items following processing of the memory instruction on the current trial; they are not critical for testing the predictions of the passive and active views of forgetting. Instead, the key predictions are based on patterns obtained at the relatively short instruction–probe SOAs. Focusing on these short SOAs, if probe RTs are longer following R instructions than following F instructions, we will be unable to distinguish between the passive view of directed forgetting (Table 1: I) and an active view in which the mechanism engaged by an F instruction is less demanding than that engaged by an R instruction (Table 1: II). If we find equivalent probe RTs following both R and F instructions (Table 1: III), this will provide weak support for an active view in which the mechanisms engaged by R and F instructions are equally demanding. (The evidence will be weak because it derives from accepting the null hypothesis of no difference between post-R and post-F probe RTs.) Finally, finding that probe RTs following F instructions are longer than those following R instructions (Table 1: IV), will provide strong evidence for an active view in which forgetting is more cognitively demanding than remembering is.

Given that the probe task required a simple detection, we expected uniformly high accuracy, such that all predictions are made with respect to RTs only.

## EXPERIMENT 1

In Experiment 1, a single detection probe was presented during each trial of the study phase. This probe could appear following either the word or the memory instruction of each trial. Whether measured relative to the onset of the word or memory instruction, detection probes were presented at one of three SOAs (1,400, 1,800, or 2,600 msec). The variable SOAs were intended to make probe onset unpredictable and to provide a time course for any post-R versus post-F differences that might have occurred. Be-

cause we were interested in illuminating the cognitive consequences of a putative attentional inhibitory mechanism initiated by an F instruction, we selected a range of SOAs over which inhibition of return (which, under some conditions, can be characterized as a spatially tagged attentional inhibition; see Taylor & Klein, 2000) is known to occur (see Samuel & Kat, 2003). Although the dependent measure of interest was postinstruction probe RT, postword (preinstruction) probes were included to ensure unpredictability of probe onset. Postword probe RTs were not of theoretical interest and were not analyzed.

A *yes–no* recognition task was used to confirm a directed forgetting effect. We used a recognition task rather than a recall task largely for pragmatic reasons. First, recognition is a valid measure of item-method directed forgetting, and robust directed forgetting effects are obtained with recognition tests when an item-method paradigm is used (e.g., Basden et al., 1994; Basden et al., 1993; see also Basden & Basden, 1998). Second, our primary goal was to evaluate probe reaction times (RTs) at several intervals relative to R and F instructions. Because we needed a sufficient number of observations within each cell of the design to obtain reliable RTs, this meant including a relatively large number of trials (60 R trials and 60 F trials). We anticipated that participants would feel helpless trying to commit 60 R words to memory for a later recall test, but that they would be motivated to do so for a subjectively easier recognition test. Finally, with this relatively large number of R and F words, we were not concerned about ceiling effects in recognition performance.

## Method

### Participants

Twenty-nine undergraduate students (20 female, 9 male) in an eligible psychology course at Dalhousie University participated in exchange for course credit. All reported normal or corrected-to-normal vision and a good understanding of the English language.

### Stimuli and Apparatus

The experiment was conducted in a small, unremarkable room with overhead lighting. Participants were tested individually in a single session lasting less than 1 h. Stimulus presentation and response collection were controlled by experimental software written

**Table 1**  
**Summary of the Predictions for the Passive and Active Accounts of Item-Method Directed Forgetting at Short and Long Instruction–Probe SOAs, Qualified by the Presence or Absence of Cumulative Rehearsal**

	Short Instruction–Probe SOAs	Long Instruction–Probe SOAs
I. Passive		
a. No cumulative rehearsal	$RT_{\text{post-R}} > RT_{\text{post-F}}$	$RT_{\text{post-R}} > RT_{\text{post-F}}$
b. Cumulative rehearsal	$RT_{\text{post-R}} > RT_{\text{post-F}}$	$RT_{\text{post-R}} = RT_{\text{post-F}}$
II. Active: Forgetting is less effortful than remembering		
a. No cumulative rehearsal	$RT_{\text{post-R}} > RT_{\text{post-F}}$	$RT_{\text{post-R}} > RT_{\text{post-F}}$
b. Cumulative rehearsal	$RT_{\text{post-R}} > RT_{\text{post-F}}$	$RT_{\text{post-R}} = RT_{\text{post-F}}$
III. Active: Forgetting is as effortful as remembering		
a. No cumulative rehearsal	$RT_{\text{post-R}} = RT_{\text{post-F}}$	$RT_{\text{post-R}} = RT_{\text{post-F}}$
b. Cumulative rehearsal	$RT_{\text{post-R}} = RT_{\text{post-F}}$	$RT_{\text{post-R}} = RT_{\text{post-F}}$
IV. Active: Forgetting is more effortful than remembering		
a. No cumulative rehearsal	$RT_{\text{post-R}} < RT_{\text{post-F}}$	$RT_{\text{post-R}} < RT_{\text{post-F}}$
b. Cumulative rehearsal	$RT_{\text{post-R}} < RT_{\text{post-F}}$	$RT_{\text{post-R}} = RT_{\text{post-F}}$

in PsyScope 5.1.2 (Cohen, MacWhinney, Flatt, & Provost, 1993) loaded on a Macintosh G4-400 computer running OS 9. Stimuli were presented on a 17-in.  $1,024 \times 768$  resolution Macintosh Studio Display color monitor, and responses were recorded via a standard Macintosh Universal Serial Bus keyboard. Participants viewed the monitor from a distance of approximately 45 cm. All words and instructions were presented against a white background in a black, 18-point Helvetica font; visual probes consisted of a black asterisk (" \* ") presented in a black, 24-point Helvetica font. The memory instructions consisted of one high-frequency (1170-Hz) tone and one low-frequency (260-Hz) tone, each presented for 400 msec via the built-in computer speakers.

The Paivio, Yuille, and Madigan Word List Generator ([www.math.yorku.ca/SCS/Online/paivio/](http://www.math.yorku.ca/SCS/Online/paivio/)) was used to create a master list of 240 randomly selected nouns. The words had an average Kučera–Francis word frequency (Kučera & Francis, 1967) of 54.63, an imagery rating of 5.26, a concreteness rating of 5.20, and a meaningfulness rating of 6.27. The words were 3–12 letters long, with a mean of 6.08 letters and 1.93 syllables. Before testing each participant, custom software was used to assign words randomly from the master list of forget ( $n = 60$ ), remember ( $n = 60$ ), and foil ( $n = 120$ ) words; this resulted in unique list compositions for each participant.

### Procedure

After giving informed consent, participants were instructed that they would view a series of study trials in which each word, presented singly, would be followed by an R or an F instruction. For half of the participants, the high-frequency tone served as the R instruction and the low-frequency tone served as the F instruction; the reverse designation was used for the other half. Participants were told that if they received an R instruction, they were to remember the word, and that if they received an F instruction, they were to forget the word. The instructions indicated that memory would be tested at the end of the experiment via recognition, but no explicit mention was made of the fact that recognition of both R and F words would be tested.

Although the memory task was described as the primary task, participants were also informed that they would need to make a speeded detection response to the onset of a visual probe on every trial. They were told that this probe could appear at one of a number of unpredictable intervals throughout the study trial.

Prior to starting the experiment proper, participants were given practice with the probe-detection task while the experimenter remained in the room. Practice trials were identical to study phase trials (described below), except that the study word was replaced with a string of five Xs, and no RT data were gathered or analyzed. At least 10 trials were presented, and the trials were repeated until the experimenter was confident that the participant understood the requirements of the probe task and until the participant reported feeling comfortable performing it. At this point the experiment was initialized, the verbal instructions given by the experimenter were reiterated on-screen, and the experimenter left the room. Participants were prompted to press the space bar when they were ready.

At the start of the experiment, participants were presented with eight trials to familiarize them with the memory instructions employed during the study phase. Each instruction familiarization trial presented a visual fixation stimulus (" + ") for 1,500 msec, followed by a memory instruction (high or low tone) for 400 msec, followed in turn by the 2,000-msec presentation of a sentence in the middle of the screen that indicated the meaning of the tone (e.g., "Remember the word."). A random half of the familiarization trials presented an R instruction, and the other half presented an F instruction.

**Study phase.** As depicted in Figure 1, each trial in the study phase began with the presentation of a centralized fixation stimulus (" + ") lasting 1,500 msec. Fixation was replaced by a word for 1,000 msec. This study word was drawn randomly without replacement from the list of 60 remember words or from the list of 60 forget words. After a delay of 2,000 msec, the corresponding (R or F) memory instruction (high or low tone) was presented for 400 msec. A single visual probe (" \* ") lasting 250 msec was presented on each trial and occurred with equal probability at postword SOAs of 1,400, 1,800, or 2,600 msec or at postinstruction SOAs of 1,400, 1,800, or 2,600 msec. Thus, this probe appeared either before or after the memory instruction, but not both, so that each trial contained a single probe. Participants responded to the detection of the probe by pressing the space bar with the index finger of the dominant hand, as quickly as possible; no feedback was given.

The duration of all study trial events summated to 7,350 msec, from the onset of fixation to the offset of the latest (i.e., 2,600-msec SOA postinstruction) possible probe; following these events was an enforced intertrial interval of 1,150 msec, during which no events were presented. Hence, the total trial duration was 8,500 msec.

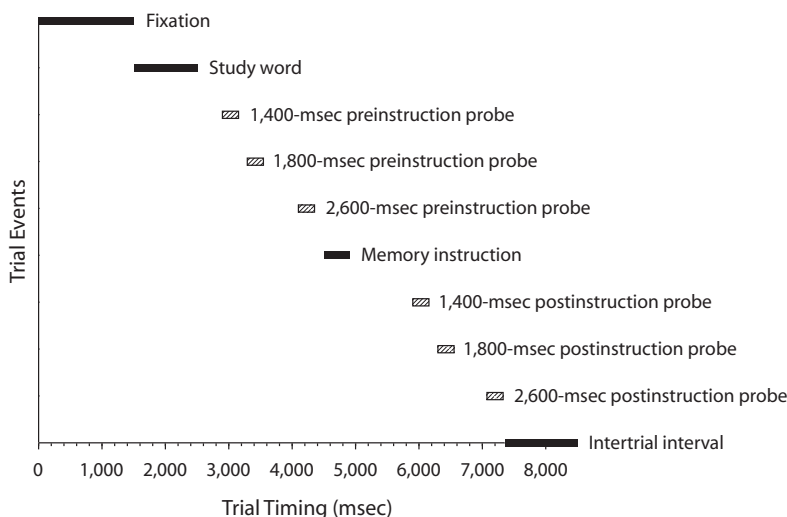


Figure 1. The timing of trial events in Experiment 1. The hatched bars represent probes; note that only one of these probe events was presented on any given trial. See the text for further details.

The design was conceived as a 2 (memory instruction: R, F)  $\times$  3 (SOA: 1,400, 1,800, 2,600 msec)  $\times$  2 (probe placement: postword, postinstruction) within-subjects factorial. The resulting 12 conditions were replicated 10 times within a block of trials, resulting in 120 study trials. However, as previously noted, the postword probes were included only to enhance the unpredictability of the probe task; they were not theoretically interesting and they were not analyzed. Thus, for the purpose of the RT analysis, the design was reconceptualized as a 2 (memory instruction: R, F)  $\times$  3 (SOA: 1,400, 1,800, 2,600 msec) design that included only postinstruction probe trials.

Twelve buffer trials (i.e., 6 at the beginning and 6 at the end) were included to minimize recency and primacy effects. Buffer trials were identical to experimental trials except that no probe RT data were gathered or analyzed. Buffer words were the same for all participants and were followed invariably by an R instruction. Recognition performance was not measured for buffer words.

**Recognition phase.** Immediately following the last trial of the study phase, the experiment advanced to the recognition phase, during which participants were presented with the 60 R and 60 F words from the study trials, intermixed randomly with 120 foil words that had not been presented at study. Detailed written instructions were presented at the top of the computer screen at all times during this phase. These instructions informed participants that they should attempt to recognize all words that were presented during study, regardless of whether an R or F memory instruction was presented. Individual words were drawn randomly without replacement and were presented in blue print one at a time (below the instructions) on the computer screen. Participants were instructed to press "Y" on the computer keyboard to indicate a word that had been presented previously (i.e., R and F words) or "N" to indicate a word that had not been presented (i.e., foil words). Responses appeared in a text box on-screen and could be changed using the backspace key or submitted using the space bar. Participants were allowed to respond to this task at their own pace. Following the recognition trials, participants were debriefed and permitted to leave.

## Results

The data from 4 participants (3 female, 1 male) were excluded due to more than 50% missed probe responses in one or more experimental conditions of the study phase. This resulted in empty cells for the RT analysis (in the case of no probe responses) or unstable RTs (because they were calculated across only a small number of trials). Statistical analyses included only the remaining 25 participants.

Although postinstruction probe RT data were the primary measure of interest, the recognition data were analyzed first to ensure that participants had attended to the memory instructions.

### Recognition Accuracy

The proportion of "yes" responses (i.e., hits to R and F words; false alarms to foils) was analyzed as a function of word type (R vs. F vs. foil) by use of a one-way repeated measures ANOVA. This analysis revealed a significant main effect of word type [ $F(2,48) = 182.69$ ,  $MS_e = 0.01$ ,  $p < .01$ ]. Planned contrasts indicated that participants correctly recognized significantly more R words ( $M = .65$ ,  $SE = .02$ ) than F words ( $M = .47$ ,  $SE = .04$ ) [ $F(1,48) = 40.77$ ,  $p < .01$ ], confirming a directed forgetting effect. An additional contrast revealed more "yes" responses to F words ( $M = .47$ ,  $SE = .04$ ) than to foil words ( $M = .11$ ,  $SE = .03$ ) [ $F(1,48) = 154.02$ ,  $p < .01$ ].

The occurrence of a significant directed forgetting effect indicated that participants attended to the memory

instructions. Therefore, the question was whether instantiating a memory instruction would influence postinstruction probe RTs in a manner consistent with the passive or active views of forgetting.

### Postinstruction Probes

Mean RTs were calculated for trials on which a correct detection response was executed to the postinstruction probe between 100 and 1,400 msec after its onset; all other trials were considered errors. In detection tasks, RTs shorter than 100 msec and longer than 1,000 msec typically are excluded as anticipations and misses, respectively (see, e.g., Lupiáñez & Milliken, 1999); in light of the simultaneous memory task, we extended the upper bound to allow for delayed responding. Although, on average, less than 1% of each participant's postinstruction probe trials were excluded by these criteria, to ensure that our cutoffs did not alter the data, we reanalyzed the postinstruction probe RTs without imposing any cutoffs; the pattern of RTs and their significance were the same as reported below. The mean postinstruction probe RTs (using the RT cutoffs) and the corresponding proportions of errors are shown in Figure 2.

An ANOVA was performed on these RT data, with memory instruction (R, F) and SOA (1,400, 1,800, 2,600 msec) as within-subjects factors. This analysis revealed significant main effects for both memory instruction [ $F(1,24) = 15.90$ ,  $MS_e = 2,908.44$ ,  $p < .01$ ] and SOA [ $F(2,48) = 7.59$ ,  $MS_e = 2,150.78$ ,  $p < .01$ ] that were qualified by a significant two-way interaction [ $F(2,48) = 4.24$ ,  $MS_e = 2,359.83$ ,  $p < .03$ ]. Contrasts indicated that, in comparison with the post-R RTs, post-F RTs were significantly longer at the 1,400-msec SOA [ $F(1,48) = 5.98$ ,  $p < .02$ ] and the 1,800-msec SOA [ $F(1,48) = 21.79$ ,  $p < .01$ ], and equivalent at the 2,600-msec SOA ( $F < 1$ ).

A memory instruction (R, F)  $\times$  SOA (1,400, 1,800, 2,600 msec) within-subjects ANOVA was also conducted

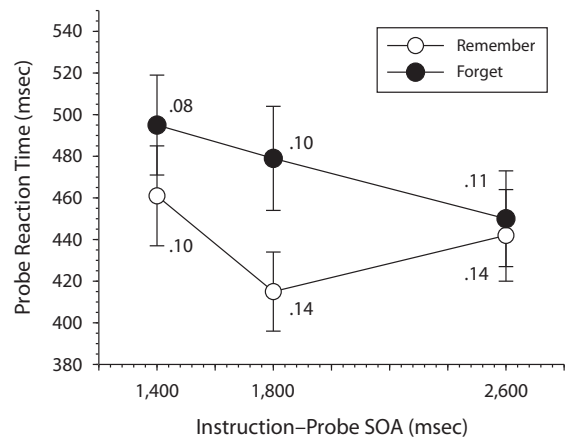


Figure 2. Mean postinstruction probe RTs (in milliseconds) in Experiment 1 as a function of memory instruction (R, F) and instruction-probe SOA (1,400, 1,800, 2,600 msec); error bars represent one standard error of the mean. Mean proportions of error are indicated for each condition.



on the error data shown in Figure 2. No effects were significant (all  $ps > .06$ ).

### Discussion

The recognition data of Experiment 1 revealed a significant directed forgetting effect. The fact that the magnitude of this effect was equivalent to that in other published studies using the item-method paradigm (e.g., Basden & Basden, 1998; these researchers reported a directed forgetting effect [R–F], roughly .20 compared with .22 in the present study) argues that performance in the memory task was not compromised by the probe-detection task.

Given that the memory instruction resulted in a typical directed forgetting effect, the question was whether the postinstruction probe RTs would fit the predictions of a passive or active view of forgetting. On this, the data were clear: The detection of post-F probes was slower than the detection of post-R probes at the two earliest SOAs (1,400 and 1,800 msec) but not at the longest SOA (2,600 msec). If we refer to Table 1, it is obvious that the obtained  $RT_{\text{post-R}} < RT_{\text{post-F}}$  pattern at the two earliest SOAs best fits the predictions of an active view in which forgetting is more effortful than remembering (Table 1: IV).

The additional finding of  $RT_{\text{post-R}} = RT_{\text{post-F}}$  at the 2,600-msec instruction–probe SOA suggests that cumulative rehearsal might have been operating at later intervals (Table 1: IVb).<sup>1</sup> However, we do not assert the latter conclusion too strongly, given that it depends on supporting a *null* hypothesis of no difference between post-R and post-F probe RTs. Moreover, the finding of no difference in RTs, although consistent with cumulative rehearsal on both R and F trials, could also reflect the completed instantiation of R and F instructions, without the engagement of any other cognitive mechanisms.

Regardless of whether cumulative rehearsal was operating at the longest SOA, it is clear that the pattern of results at the two shortest SOAs contradicts the passive variant of the selective rehearsal account: Not only is forgetting cognitively demanding, it is cognitively *more* demanding than remembering, at least in the initial periods following the instruction. This conclusion is consistent with the attentional inhibition hypothesis (see Zacks et al., 1996).

To further explore these findings, in Experiment 2 we replicated and extended the methods of Experiment 1.

### EXPERIMENT 2

In Experiment 2, we repeated the basic methods of Experiment 1, but with four key changes. First, the memory instruction was changed from an auditory to a visual stimulus. This was intended to map onto the larger literature that has tended to use visual rather than auditory instructions. Second, catch trials on which no probe was presented were also included to allow the measurement of probe false alarms. We did not expect either of these changes to alter our findings.

Third, we wanted to learn whether an explicit instruction to refrain from cumulative rehearsal would alter the pattern of post-R and post-F probe RTs. In Experiment 1,

RTs at the 2,600-msec postinstruction SOA were statistically equivalent following R and F instructions. We suggested that participants might have completed the instantiation of the memory instruction by this time and used the remaining trial duration to cumulatively rehearse R words from preceding trials. However, it was also possible that cumulative rehearsal might have occurred at earlier intervals and contributed to the  $RT_{\text{post-R}} < RT_{\text{post-F}}$  pattern that was observed. Consider that if participants initiated cumulative rehearsal immediately after being given an F instruction, the cognitively effortful retrieval of R words from preceding trials might have contributed to—or, indeed, even accounted for—the longer post-F versus post-R probe RTs at the earliest postinstruction probe SOAs. Thus, in Experiment 2, we asked participants to refrain from the cumulative rehearsal of R words and asked them to try to treat each trial as a discrete event. We did not require participants to rehearse aloud (cf. Golding et al., 1996; Rundus & Atkinson, 1970), on the grounds that doing so would introduce a third task demand to an already difficult task and/or potentially affect their rehearsal strategy (see Fischler, Rundus, & Atkinson, 1970; Kellas, McCauley, & McFarland, 1975a, 1975b). If the observed  $RT_{\text{post-R}} < RT_{\text{post-F}}$  pattern at the earliest SOAs in Experiment 1 was due to cognitive load differences for R and F instructions (rather than differences in the retrieval and cumulative rehearsal of R items from preceding trials), we expected this pattern to replicate even under instructions to refrain from cumulative rehearsal.

Finally, recognizing that an instruction to refrain from cumulative rehearsal is a weak manipulation, we included no-word control trials in which a string of Xs was presented in place of a word. We included this condition in an attempt to subtract out any effects that cumulative rehearsal might have on R and F trials. Our reasoning was based on the view that the memory instruction would have no relevance to an X-string. Thus, if participants were unable to heed our instruction to avoid cumulative rehearsal, we expected that they would engage such a strategy on these control trials (either following the presentation of an irrelevant R or F instruction on an X-string trial or immediately following the X-string). To the extent that this was true, the intention was to subtract, on a subject-by-subject basis, probe RTs on no-word (X-string) control trials from probe RTs on word trials. We performed this subtraction in two ways: We subtracted the postinstruction probe RTs on X-string trials from corresponding postinstruction probe RTs on R or F word trials; we also subtracted preinstruction (post-X-string) probe RTs on control trials from the postinstruction probe RTs on word trials. To the extent that the  $RT_{\text{post-R}} < RT_{\text{post-F}}$  pattern observed at the earliest SOAs in Experiment 1 was due to the relative cognitive load associated with R and F instructions per se (rather than to differences in the retrieval and cumulative rehearsal of preceding R words), we expected this pattern to survive both subtraction methods.

To reiterate, our strongest predictions were for relatively short instruction–probe SOAs, for which it is possible to distinguish most clearly between passive and active accounts of forgetting (see Table 1). Although Experiment 2

included manipulations intended to eliminate and/or to control for cumulative rehearsal, we did not expect these manipulations to alter our finding of longer post-F RTs than post-R RTs at the shortest SOAs (i.e., because the shortest RTs were not presumed to include effects of cumulative rehearsal). If they did, we would be inclined to believe that memory retrieval subserving cumulative rehearsal was responsible for the longer post-F differences than post-R differences that we observed in Experiment 1. However, if post-F probe RTs continued to be longer than post-R probe RTs at the shortest SOAs in Experiment 2, we would be more confident in our conclusion that forgetting is, in fact, an active process.

## Method

### Participants

Thirty-eight undergraduate students (25 female, 13 male) enrolled in an eligible psychology class at Dalhousie University were run in this experiment in exchange for course credit. Participants reported normal or corrected-to-normal vision and a good understanding of the English language.

### Stimuli and Apparatus

The stimuli and apparatus were identical to those of Experiment 1, with the following exceptions: In Experiment 2, memory instructions consisted of yellow- or pink-filled circles outlined in a 4-point black border, 2.2 cm in diameter. To increase alertness (see Bertelson & Tisseyre, 1969), fixation onset was accompanied by a 400-msec, 1170-Hz tone.

To accommodate the addition of X-string (control) and catch trials in Experiment 2, a new list of 256 nouns was created using the Paivio, Yuille, and Madigan Word List Generator. The words on this list had an average Kučera–Francis word frequency of 56.66, imagery rating of 5.44, concreteness rating of 5.33, and meaningfulness rating of 6.27. The words were 3–8 letters long, with an average of 5.46 letters and 1.64 syllables. As in Experiment 1, the master word list was divided randomly into forget ( $n = 64$ ), remember ( $n = 64$ ), and foil ( $n = 128$ ; used during recognition) lists to produce unique compositions for each participant.

### Procedure

The procedure for this study was the same as that for Experiment 1, except for the following: In the instructions, participants were informed that on some trials (i.e., control trials), a string of Xs would appear instead of a word and that the memory instruction had no meaning for such trials. They were also told that probes would not occur on every trial (i.e., catch trials). To minimize the influence of cumulative rehearsal on probe RTs, participants were presented with a verbal description of cumulative rehearsal and were asked explicitly to refrain from engaging in this strategy; instead they were asked to focus on each trial as a discrete task, and therefore not to rehearse any study words that preceded the current trial. The memory instructions used during the instruction familiarization phase and the study phase were 400-msec presentations of pink/yellow circles in the center of the computer screen. For half of the participants, the pink circle served as an R instruction and the yellow circle served as an F instruction; this designation was reversed for the other half.

**Study phase.** Participants were presented with 128 word trials, half drawn randomly without replacement from the remember list and half drawn randomly without replacement from the forget list. Control trials were identical to word trials, except that instead of a study word, a string of five Xs (i.e., “XXXXX”) was presented. Of the 128 X-string control trials, half contained an R instruction and half contained an F instruction (even though such instructions were meaningless in this context). Word and control trials were randomized and were intermixed into a single block.

As shown in Figure 3, each trial began with the presentation of a centralized fixation stimulus (i.e., “+”) lasting 1,500 msec, accompanied by a 400-msec tone intended to warn participants of the upcoming trial events. The fixation was replaced by a word or X-string, depending on whether the trial was a word or control trial. This stimulus was presented for 1,000 msec and was replaced 2,400 msec later by the memory instruction (i.e., pink or yellow circle), which lasted 400 msec. Note that the word–memory instruction interval was 400 msec longer in Experiment 2 than in Experiment 1; this ensured enough time for a response to be executed to the 2,600-msec postword probe before the onset of the subsequent memory instruction.<sup>2</sup> On probe trials, a 250-msec visual probe (“\*”) was presented with equal probability at one of six SOAs (1,400, 1,800, or 2,600 msec in relation to the onset of the study word/X-string or at 1,400, 1,800, or 2,600 msec in relation to the onset of the memory instruction). On 64 of the trials (catch trials), no visual probe was

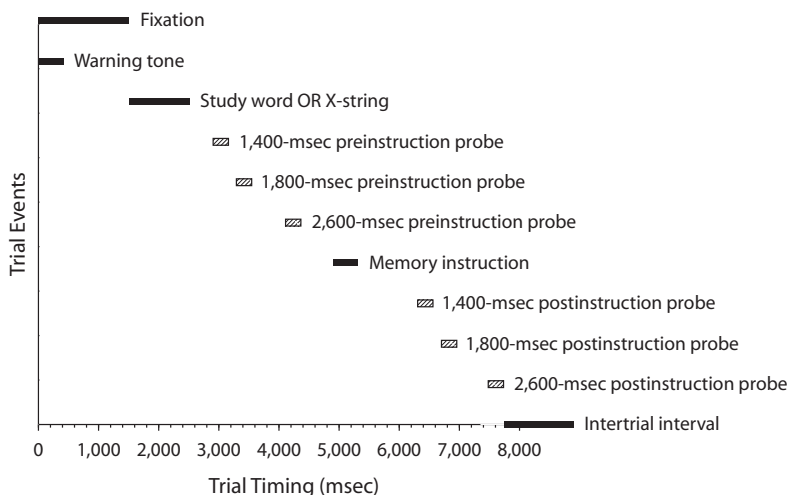


Figure 3. The timing of trial events in Experiment 2. The hatched bars represent probes; note that only one of these probe events was presented on any given trial. See the text for further details.

presented. Due to the 400-msec increase between the target word and memory instruction, each trial now lasted 8,900 msec.

The design for probe trials was conceived as a 2 (memory instruction: R, F)  $\times$  3 (SOA: 1,400, 1,800, 2,600 msec)  $\times$  2 (probe placement: postword, postinstruction)  $\times$  2 (trial type: word trial, control trial) within-subjects factorial design. The ensuing 24 conditions were replicated 8 times within a trial block, for a total of 192 probe trials. Catch trials were presented equally often across all levels of instruction and trial type factors and were replicated 16 times within a trial block, for a total of 64 catch trials. This resulted in a combined count of 256 study trials. A total of 6 buffer trials (3 at the beginning, 3 at the end) were also presented, as in Experiment 1.

For the purpose of analysis, the experimental trials were reconceptualized as a 2 (memory instruction: R, F)  $\times$  3 (SOA: 1,400, 1,800, 2,600 msec) design for which postinstruction probe RTs were the dependent measure of interest.

**Recognition phase.** The recognition phase randomly presented the 64 R and 64 F study words, along with 128 foil words that had not been presented during the study trials.<sup>3</sup>

## Results

Data from 2 female participants were excluded from statistical analysis due to the exhibition of behavior incongruent with focused task performance (e.g., using a cell phone during the experimental trials). Statistical analyses included only the remaining 36 participants.

As was done for Experiment 1, a manipulation check of the recognition data was performed first, to confirm a directed forgetting effect. Following this, postinstruction probe RTs and proportions of error were examined for word trials. This was followed by consideration of postinstruction probe RTs and proportions of error for control trials and by two different subtractions of word and control trial probe RTs. Finally, catch trial false alarms were examined for word and control trials.

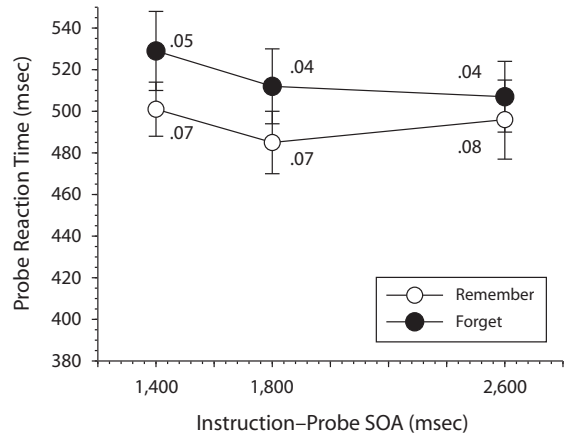
### Recognition Accuracy

The proportion of "yes" responses (i.e., hits to R and F words; false alarms to foils) was analyzed as a function of word type (R vs. F vs. foil) in a one-way repeated measures ANOVA. The results indicated a significant main effect of word type [ $F(2,70) = 181.23$ ,  $MS_e = 0.02$ ,  $p < .01$ ]. Confirming a directed forgetting effect, planned comparisons found that participants correctly recognized more R words ( $M = .66$ ,  $SE = .03$ ) than F words ( $M = .53$ ,  $SE = .03$ ) [ $F(1,70) = 19.74$ ,  $p < .01$ ]. They also recognized significantly more F words than they did foils ( $M = .13$ ,  $SE = .02$ ) [ $F(1,70) = 190.75$ ,  $p < .01$ ].

To determine whether the request to refrain from cumulative rehearsal had an impact on the magnitude of the directed forgetting effect, we compared the directed forgetting effect (i.e., R–F recognition hits) in a between-subjects analysis, with experiment (Experiment 1, Experiment 2) as the only factor. This analysis revealed no significant difference ( $M_{\text{Experiment 1}} = 0.18$ ,  $M_{\text{Experiment 2}} = 0.13$ ) [ $F(1,59) = 2.96$ ,  $MS_e = 0.01$ ,  $p > .09$ ].

### Postinstruction Probes

Mean RTs were calculated for trials on which a correct response was executed within 100 msec and 1,400 msec of postinstruction probe onset, and are shown in Figure 4; all other trials were considered errors. On average, these



**Figure 4.** Mean postinstruction probe RTs (in milliseconds) in Experiment 2 as a function of memory instruction (R, F) and instruction-probe SOA (1,400, 1,800, 2,600 msec); error bars represent one standard error of the mean. Mean proportions of error are indicated for each condition.

criteria excluded less than 1% of each participant's postinstruction probe trials; as with Experiment 1, eliminating these cutoffs did not affect the overall pattern of RTs in comparison with what is reported below. The mean RT data (using cutoffs) and the associated errors are shown in Figure 4.

Postinstruction probe RTs were analyzed in a memory instruction (R, F)  $\times$  SOA (1,400, 1,800, 2,600 msec) within-subjects ANOVA. This analysis revealed significant main effects of memory instruction [ $F(1,35) = 9.63$ ,  $MS_e = 2,709.45$ ,  $p < .01$ ] and SOA [ $F(2,70) = 3.15$ ,  $MS_e = 1,735.39$ ,  $p < .05$ ]. The memory instruction  $\times$  SOA interaction failed to reach significance ( $F < 1$ ). Nevertheless, planned comparisons were used to compare probe RTs on R and F trials at each postinstruction SOA (see Keppel, 1982). Relative to post-R probe RTs, post-F probe RTs were longer at both the 1,400-msec SOA [ $F(1,70) = 5.08$ ,  $p < .03$ ] and 1,800-msec SOA [ $F(1,70) = 4.79$ ,  $p < .04$ ], but not significantly different at the 2,600-msec SOA ( $F < 1$ ).

An analogous ANOVA conducted on the proportions of error shown in Figure 4 revealed only a main effect of memory instruction [ $F(1,35) = 6.77$ ,  $MS_e = 0.01$ ,  $p < .02$ ], with .03 fewer errors made to post-F than to post-R probes (all other  $ps > .10$ ). Importantly, the main effect of memory instruction in the error data was not indicative of a speed-accuracy trade-off in the postinstruction probe RT data. Whereas a sign test revealed that a significant number of participants were slower to respond to post-F than to post-R probes ( $n+ = 25$ ,  $n- = 11$ ,  $p < .03$ ), a nonsignificant number of participants committed fewer errors post-F than post-R ( $n+ = 19$ ,  $n- = 5$ ,  $p > .42$ ). Moreover, in those participants who made fewer errors post-F than post-R, the F–R postinstruction probe RT difference tended to be smaller (9 msec) than for those who did not show this error pattern (35 msec) [ $F(1,34) = 3.638$ ,  $MS_e = 1,633.843$ ,  $p = .065$ ].



**Table 2**  
**Mean Post-X-String and Postinstruction Probe Reaction Times**  
**(RTs, in Milliseconds) and Associated Proportion of Errors**  
**for Control Trials in Experiment 2 As a Function of Memory**  
**Instruction and Instruction-Probe SOA**

	Memory Instruction							
	Remember				Forget			
	RT		Errors		RT		Errors	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Post-X-String SOA								
1,400 msec	583	14	.09	.02	572	12	.06	.01
1,800 msec	547	12	.06	.02	535	12	.04	.01
2,600 msec	519	14	.04	.02	511	13	.05	.02
Post-Instruction SOA								
1,400 msec	507	19	.06	.02	503	14	.05	.02
1,800 msec	483	18	.04	.01	500	13	.03	.02
2,600 msec	481	14	.05	.02	492	13	.04	.01

### Control Trial Probes

Because they would be used in subsequent subtractions from probe RTs on word trials, the probe RTs on control trials were analyzed. One analysis examined the probe RTs that occurred following the X-string; another examined probe RTs following the irrelevant memory instruction. These data are shown in Table 2, along with associated errors.

An analysis of the probe RTs that followed the irrelevant memory instruction revealed no significant main effects or interactions [SOA:  $F(2,70) = 2.93$ ,  $MS_e = 2,231.03$ ,  $p = .06$ ; all other  $ps > .10$ ]. An analysis of the probe RTs that followed the presentation of the X-string revealed a main effect only of SOA [ $F(2,70) = 33.362$ ,  $MS_e = 2,160.90$ ,  $p < .01$ ]; no other effects or interactions reached significance (all other  $ps > .10$ ).

Analogous ANOVAs conducted on the corresponding error data for postinstruction and post-X-string probes revealed no significant main effects or interactions (all  $ps > .09$ ).

### Word-Control Trial Probe RTs

To subtract out any effects of cumulative rehearsal that might have occurred despite our instructions to refrain from such rehearsal, postinstruction probe RTs on control trials were subtracted from those on word trials. These difference scores are shown in Table 3. To this point, we have presumed that longer RTs following F instructions versus

R instructions at short instruction-probe SOAs on word trials means that forgetting is more cognitively demanding than remembering, which is consistent with the active view of directed forgetting. If, instead, longer RTs following F instructions are due to the effortful retrieval and cumulative rehearsal of R items (from previous trials), then this subtraction should eliminate the post-F versus post-R difference in RTs (on word trials). Planned comparisons revealed that the larger difference scores post-F than post-R continued to be significant at the 1,400-msec SOA [ $F(1,70) = 4.51$ ,  $MS_e = 3,931.93$ ,  $p < .04$ ], but not at the 1,800-msec SOA ( $F < 1$ ) or at the 2,600-msec SOA ( $F < 1$ ).

Our second subtraction method is predicated on the following: If participants engaged in an effortful retrieval and cumulative rehearsal of previous R words during the control trials, they might have engaged this strategy after the presentation of the X-string, rather than waiting until after the presentation of the (irrelevant) memory instruction that followed the X-string. Thus, we also subtracted post-X-string probe RTs on control trials from the postinstruction probe RTs on word trials. An analysis of the resulting difference scores, also shown in Table 3, revealed that post-F probe RTs were significantly longer than post-R probe RTs at the 1,400-msec SOA [ $F(1,70) = 5.83$ ,  $MS_e = 4,621.37$ ,  $p < .02$ ] and 1,800-msec SOA [ $F(1,70) = 6.04$ ,  $MS_e = 4,621.37$ ,  $p < .02$ ] and were statistically equivalent to post-R trials at the 2,600-msec SOA [ $F(1,70) = 1.39$ ,  $MS_e = 4,621.37$ ,  $p > .24$ ].

### No-Probe Catch Trials

False alarms on no-probe (catch) trials were analyzed separately for word and control trials as a function of memory instruction (i.e., R vs. F) via a one-way within-subjects ANOVA. For word trials, there was a significant main effect of memory instruction [ $F(1,35) = 25.46$ ,  $MS_e = 0.31$ ,  $p < .01$ ], indicating a significantly higher proportion of false alarms following R instructions ( $M = .05$ ,  $SE = .01$ ) than following F instructions ( $M = .01$ ,  $SE = .01$ ). For control trials, the difference in false alarms following R instructions ( $M = .03$ ,  $SE = .01$ ) and F instructions ( $M = .02$ ,  $SE = .01$ ) was not significant [ $F(1,35) = 2.94$ ,  $MS_e = 0.38$ ,  $p > .05$ ].

### Discussion

Experiment 2 broadly replicated the main findings of Experiment 1, despite several changes to the methods,

**Table 3**  
**Mean Word-Control Probe Reaction Time (RT, in Milliseconds)**  
**Difference Scores in Experiment 2, Using Post-X-String or**  
**Postinstruction Control Probes As the Subtractive Baseline, As a**  
**Function of Memory Instruction and Instruction-Probe SOA**

Instruction-Probe SOA	Subtractive Baseline							
	Post-X-String				Postinstruction			
	Remember		Forget		Remember		Forget	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
1,400 msec	-82	15	-43	15	-6	14	26	12
1,800 msec	-62	12	-22	13	2	12	13	11
2,600 msec	-23	16	-4	14	15	12	15	10

including the following: (1) the use of visual memory instructions, (2) an instruction to refrain from cumulative rehearsal, (3) the inclusion of no-probe catch trials, and (4) the inclusion of no-word (X-string) control trials. As in Experiment 1, the word trials revealed that probe RTs following F instructions were longer than those following R instructions at the 1,400- and 1,800-msec SOAs. This finding is consistent with the predictions of the active view of item-method directed forgetting in which forgetting is more effortful than remembering (Table 1: IV); it directly contradicts the predictions of the passive view (Table 1: I).

Efforts were made in Experiment 2 to eliminate and/or control for any effects of the retrieval and cumulative rehearsal of R words across trials. One attempt at such control was our request that participants refrain from cumulative rehearsal. To the extent that participants were capable of heeding this request, it had no effect on the overall recognition of R words ( $M = .65$  in Experiment 1 and  $M = .66$  in Experiment 2;  $F < 1$ ) or F words [ $M = .47$  in Experiment 1,  $M = .53$  in Experiment 2;  $F(1,59) = 2.00$ ,  $MS_e = 0.06$ ,  $p > .15$ ]. At first blush, the equivalent R recognition in Experiments 1 and 2 might seem surprising. However, requesting that participants refrain from across-trial cumulative rehearsal is not the same as requesting that they refrain from elaborative rehearsal. Indeed, we fully expected them to rehearse each R word elaboratively, but to do so only in the context of that particular R word trial. To the extent that they refrained from across-trial cumulative rehearsal, the equivalence of Experiment 2 recognition data to those of Experiment 1 suggests that cumulative rehearsal offers little additional benefit to the elaborative rehearsal that R words otherwise receive within individual trials. This is not to say that cumulative rehearsal is *never* helpful in such a paradigm, but there are two primary reasons why it may not have been beneficial in the present experiments: (1) Participants were expected to retain a large number of R words (60 in Experiment 1; 64 in Experiment 2), and (2) study phase trials permitted only a few seconds of rehearsal time between successive word presentations (word-word intervals of 8.5 sec in Experiment 1 and 8.9 sec in Experiment 2), during which time the probe task was performed. Because of these factors, we suspect that only a small number of R words would have received any degree of across-trial cumulative rehearsal anyway (see note 1).

Regardless, we are aware that our instruction to refrain from across-trial cumulative rehearsal constituted a weak manipulation of this behavior. For this reason, we included the X-string control trials. Presuming that a memory instruction applied to an X-string, or to the presentation of the X-string itself, would likewise encourage the use of cumulative rehearsal (to the extent that this strategy could not be avoided, despite task instructions), Experiment 2 used probe RTs on control trials to subtract out any presumed effects of cumulative rehearsal on word trials. Even after these subtractive measures were applied, post-F probe RTs continued to be longer than post-R probe RTs at the earliest SOA (the data were less robust at the 1,800-msec SOA). This suggests that longer post-F

than post-R probe RTs at this early SOA were most likely due to greater cognitive demands associated with forgetting than with remembering (Table 1: IV).

Interestingly, the results of Experiment 2 also revealed that participants committed fewer false alarms during F-instructed than during R-instructed catch trials, but only when a word was presented (therefore making the instruction task-relevant). This suggests that the instantiation of an F instruction may limit spurious responses and/or that the instantiation of an R instruction may encourage them.

## GENERAL DISCUSSION

The present experiments used a variant of the item-method directed forgetting paradigm to evaluate competing views of forgetting as due to passive decay versus an active cognitive process. Participants were presented with a series of words, each followed by an instruction to remember or to forget. For both experiments, during a subsequent memory task, recognition was greater for R words than it was for F words, indicating a directed forgetting effect. To determine whether this R-F difference in recognition performance was mediated by passive decay, as suggested by the standard interpretation of the selective rehearsal account (e.g., Basden et al., 1993) or by an active cognitive mechanism such as attentional inhibition (e.g., Zacks et al., 1996), we embedded a detection probe in the study phase. Postinstruction probe RTs were longer following F instructions than following R instructions at the shortest SOAs in both experiments, providing an index of relative cognitive load (see Kahneman, 1973).

These findings are incongruent with a passive decay version of the selective rehearsal account (see Table 1). If an R instruction led to elaborative rehearsal of R items and an F instruction led words to be “dropped” from the rehearsal set (e.g., Basden et al., 1993), longer RTs would be expected following R instructions than following F instructions. Instead, our findings best fit the predictions of an active model of forgetting, such as the attentional inhibition model proposed by Zacks et al. (1996). Zacks et al. argued that an F instruction engages an inhibitory mechanism that limits further processing of irrelevant information (F words) in order to conserve cognitive resources for the processing of relevant information (R words). The operation of such a mechanism easily accounts for longer post-F instruction probe RTs than post-R instruction probe RTs.

A steadfast proponent of the standard selective rehearsal account might argue that we have been too hasty in disregarding a passive decay view of forgetting: It is possible that participants passively drop the F item from memory and, in doing so, free up cognitive resources to retrieve and cumulatively rehearse R words from preceding trials; according to this argument, it is the effortful retrieval of preceding R words that leads to longer probe RTs post-F than post-R. Furthermore, the equivalent post-F and post-R probe RTs at the longest SOA could be accounted for by proposing that the memory search in the F condition terminated at some point between 1,800

and 2,600 msec after the memory instruction, thereby eliminating the RT disadvantage for post-F versus post-R probe detection.

Although this passive decay hypothesis is consistent with the present results, the problem with this "rescue" of this hypothesis is that it must further presume that there is no memory retrieval on R trials and/or that any such retrieval is not as effortful as that which occurs on F trials. In a typical item-method directed forgetting task, it is likely that participants perform some kind of relational processing, linking R items together across trials. This suggestion is corroborated by informal postexperimental reports by our participants and likely invalidates the critical assumption that participants engage in retrieval of previous R words only following an F instruction. In the process of elaborately rehearsing R words on a given trial and relating them to previous R words, participants are likely to retrieve the representation of those previous R words. As such, a cognitively effortful search for and retrieval of those words should be required on R trials, just as was postulated for the F condition.

Although we believe that an active view provides the most parsimonious explanation for the observed  $RT_{\text{post-R}} < RT_{\text{post-F}}$  pattern of results, we accept that cumulative rehearsal of preceding R words *may* occur on at least some forget trials. Nevertheless, two aspects of Experiment 2 argue that cumulative rehearsal cannot provide a *full* explanation of the  $RT_{\text{post-R}} < RT_{\text{post-F}}$  pattern that occurred at the early SOAs. First, in Experiment 2, participants were explicitly asked to refrain from cumulative rehearsal; informal, postexperimental participant reports indicated compliance. To the extent that they were able to limit or prevent their use of this strategy, there is no evidence that it affected the pattern of probe detection RTs: In Experiment 2, as in Experiment 1, postinstruction probe RTs were longer following F instructions than following R instructions at the earliest SOAs. Second, subtracting the probe RTs on control trials from probe RTs on word trials should have eliminated any remaining effects of cumulative rehearsal, yet this subtraction did not eliminate the difference between F and R trials at the 1,400-msec SOA (the post-F vs. post-R difference at the 1,800-msec SOA was less robust).

Despite the persuasive nature of the probe RT data, the analyses presented above have not yet connected probe RT differences directly to subsequent memory performance. To maximize statistical power, post-F and post-R RTs were pooled across Experiments 1 and 2, were collapsed across SOA, and were analyzed separately as a function of outcome on the subsequent recognition test (remembered vs. forgotten). This analysis required probe RTs for each combination of memory instruction and recognition outcome (R remembered, R forgotten; F remembered, F forgotten); this requirement was met in the data contributed by 59 participants. The analysis of F trials indicated that participants tended to be slower to respond to probes that followed F words that were later forgotten ( $M_{F\text{-forgotten}} = 500$  msec,  $SE_{F\text{-forgotten}} = 15$  msec) than they were to those that were later remembered ( $M_{F\text{-remembered}} = 485$  msec,  $SE_{F\text{-remembered}} = 13$  msec) [ $F(1,58) = 3.60$ ,

$MS_e = 6,398.5$ ,  $p = .06$ ]. There was no such tendency for R words ( $M_{R\text{-forgotten}} = 468$  msec,  $SE_{R\text{-forgotten}} = 15$  msec;  $M_{R\text{-remembered}} = 469$  msec,  $SE_{R\text{-remembered}} = 12$  msec) ( $F < 1$ ). This analysis suggests that successfully forgetting irrelevant information (F words) is associated with an increase in postinstruction probe RTs, and it provides behavioral evidence that whatever happened following an F instruction affected subsequent attempts to recognize the to-be-forgotten word (for similar evidence from functional neuroimaging, see Reber et al., 2002; Wylie, Foxe, & Taylor, 2008). On the basis of our pattern of longer RTs post-F than post-R, we argue that what happened following an F instruction was an active cognitive process associated with the suppression of the representation of the F word, such as was predicted by the attentional inhibition hypothesis.

Critically, the inhibitory mechanism implicated in item-method directed forgetting is distinguishable from that which sometimes has been presumed to operate in list-method directed forgetting (e.g., Geiselman et al., 1983) or the think/no-think (e.g., Anderson & Green, 2001; Anderson et al., 2004) paradigms. Whereas the latter two paradigms have been associated with inhibition operating at *retrieval*, attentional inhibition in the item-method paradigm is presumed to operate primarily at *encoding* (although see Ullsperger, Mecklinger, & Müller, 2000). In this respect, the attentional inhibition account is not necessarily at odds with a selective rehearsal account per se; it is at odds only with a *passive* view of selective rehearsal in which F words are passively "dropped" from the rehearsal set. Indeed, attentional withdrawal from F items and subsequent inhibition from returning to their representation (see Zacks et al., 1996) may provide a mechanism by which limited-capacity resources are reallocated from irrelevant (F) to relevant (R) items so that the latter can be committed to memory.

Recent research provides additional converging evidence for attentional involvement in item-method directed forgetting. Taylor (2005; see also Taylor & Fawcett, 2008) demonstrated larger magnitude inhibition of return following peripherally presented F words than following peripherally presented R words. This finding is consistent with the view that visuospatial attention withdraws more readily from the spatial representation of peripherally presented F words than from that of peripherally presented R words. For the case in which words are presented centrally rather than peripherally, thereby eliminating a role for visuospatial attention, Hourihan and Taylor (2006) have argued that executive attentional control mechanisms are engaged by an F instruction to cease the commitment of words to memory. In this way, an instruction to forget in an item-method task is analogous to a stop signal in a behavioral countermanding paradigm (for a review, see Logan, 1994). Indeed, event-related functional imaging reveals that the same areas in the inferior frontal gyrus that are involved in the cessation of overt responses (e.g., Aron, Fletcher, Bullmore, Sahakian, & Robbins, 2003; Overtom et al., 2002; Pliszka, Liotti, & Woldorff, 2000; Rieger, Gauggel, & Burmeister, 2003) are more active during the study of an F word that

is later successfully forgotten than during the study of an F word that is accidentally remembered (Wyllie et al., 2008; see also Reber et al., 2002).

Although our study is not the first to quantify the cognitive demands of intentional forgetting, it is the first long-term memory study to expressly do so by using a conventional item-method task. Past attempts have often dealt with short-term memory variants of the item-method paradigm (e.g., Roediger & Crowder, 1972) or have employed a list-method paradigm (e.g., Johnson, 1971; Martin & Kelly, 1974). Most relevantly, Roediger and Crowder found that in a Brown–Peterson short-term memory distractor task, participants produce more numbers following an instruction to forget than they did following an instruction to remember (i.e., they counted *faster* following F than following R instructions). Although their finding initially may appear incongruous with our own, two things must be considered. First, Roediger and Crowder employed a distractor task that extended between 4,500 and 18,500 msec following concurrent trigram–memory instruction onset, well beyond the postinstruction interval during which we observed longer post-F than post-R probe RTs. Second, attentional withdrawal from the F word would, in fact, predict faster backward counting in the Brown–Peterson task: By ejecting the F word from working memory, the cognitive resources previously dedicated to maintaining its representation are released for use during the distractor task. In this respect, it seems likely that the slower post-F versus post-R RTs in our task reflect the cognitively demanding act of instantiating the F instruction (viz., attentional inhibition) and that the Roediger and Crowder findings reflect the benefit of having done so: Attention is removed from the irrelevant F items, thereby freeing cognitive resources for other tasks.

In light of the converging evidence presented above, the present experiments are critical. These experiments represent the first attempt to explicitly test the predictions of passive and active views of directed forgetting using a behavioral paradigm in a nonclinical sample of young adults. Our postinstruction probe RTs suggest not only that forgetting is an active cognitive process, but that it is, at least initially, cognitively *more* demanding than is remembering. Interpreted in relation to current views of item-method directed forgetting, these findings contradict the passive variant of the selective rehearsal account and provide strong support for an active cognitive model, such as the attentional inhibition view proposed by Zacks et al. (1996).

#### AUTHOR NOTE

This research was funded by a grant to T.L.T. from the Natural Sciences Engineering and Research Council of Canada (NSERC); J.M.F. was supported by an NSERC Canada Graduate Scholarship and a Killam Predoctoral Scholarship. We thank Carl Helmick for writing the custom software that was used to distribute words to lists. Thanks also to Alison Moss, Ray Klein, Vincent LoLordo, and three anonymous reviewers for feedback on earlier versions of this manuscript. Address correspondence regarding this article to J. M. Fawcett, Department of Psychology, Dalhousie University, 1355 Oxford Street, Halifax, NS, B3H 4J1 Canada (e-mail: jmfawcett@dal.ca).

#### REFERENCES

- ANDERSON, M. C., & GREEN, C. (2001). Suppressing unwanted memories by executive control. *Nature*, **410**, 366–369.
- ANDERSON, M. C., OCHSNER, K. N., KUHL, B., COOPER, J., ROBERTSON, E., GABRIELI, S. W., ET AL. (2004). Neural systems underlying the suppression of unwanted memories. *Science*, **303**, 232–235.
- ARON, A. R., FLETCHER, P. C., BULLMORE, E. T., SAHAKIAN, B. J., & ROBBINS, T. W. (2003). Stop-signal inhibition disrupted by damage to right inferior frontal gyrus in humans. *Nature Neuroscience*, **6**, 115–116.
- BASDEN, B. H., & BASDEN, D. R. (1998). Directed forgetting: A contrast of methods and interpretations. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 139–172). Mahwah, NJ: Erlbaum.
- BASDEN, B. H., BASDEN, D. R., COE, W. C., DECKER, S., & CRUTCHER, K. (1994). Retrieval inhibition in directed forgetting and posthypnotic amnesia. *International Journal of Clinical & Experimental Hypnosis*, **42**, 184–203.
- BASDEN, B. H., BASDEN, D. R., & GARGANO, G. J. (1993). Directed forgetting in implicit and explicit memory tests: A comparison of methods. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 603–616.
- BERTELSON, P., & TISSEYRE, F. (1969). The time-course of preparation: Confirmatory results with visual and auditory warning signals. *Acta Psychologica*, **30**, 145–154.
- BJORK, R. A. (1970). Positive forgetting: The noninterference of items intentionally forgotten. *Journal of Verbal Learning & Verbal Behavior*, **9**, 255–268.
- BJORK, R. A. (1972). Theoretical implications of directed forgetting. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 217–235). Washington, DC: Winston.
- 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.
- COHEN, J. D., MACWHINNEY, B., FLATT, M., & PROVOST, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavior Research Methods, Instruments, & Computers*, **25**, 257–271.
- CONWAY, M. A., & FTHENAKI, A. (2003). Disruption of inhibitory control of memory following lesions to the frontal and temporal lobes. *Cortex*, **39**, 667–686.
- CONWAY, M. A., HARRIES, K., NOYES, J., RACSMA'NY, M., & FRANKISH, C. R. (2000). The disruption and dissolution of directed forgetting: Inhibitory control of memory. *Journal of Memory & Language*, **43**, 409–430.
- CRAIK, F. I. [M.], & LOCKHART, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning & Verbal Behavior*, **11**, 671–684.
- ELMES, D. (1969). Cueing to forget in short-term memory. *Journal of Experimental Psychology*, **80**, 561–562.
- FISCHLER, I., RUNDUS, D., & ATKINSON, R. C. (1970). Effects of overt rehearsal procedures on free recall. *Psychonomic Science*, **19**, 249–250.
- GEISELMAN, R. E., BJORK, R. A., & FISHMAN, D. L. (1983). Disrupted retrieval in directed forgetting: Evidence for retrieval inhibition. *Journal of Experimental Psychology: General*, **112**, 58–72.
- GOLDING, J. M., ROPER, K. L., & HAUSETL, J. (1996). To forget or not to forget: The effect of probability of test on directed forgetting. *Quarterly Journal of Experimental Psychology*, **49A**, 326–340.
- HASHER, L., & ZACKS, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22, pp. 193–223). San Diego: Academic Press.
- HENKEL, L. A., JOHNSON, M. K., & DE LEONARDIS, D. M. (1998). Aging and source monitoring: Cognitive processes and neuropsychological correlates. *Journal of Experimental Psychology: General*, **127**, 251–268.
- HOURLIHAN, K. L., & TAYLOR, T. L. (2006). Cease remembering: Executive control processes in directed forgetting. *Journal of Experimental Psychology: Human Perception & Performance*, **32**, 1354–1365.



- JAMES, W. (1950). *The principles of psychology* (Vol. 2). New York: Dover. (Original work published 1890)
- JOHNSON, D. A. (1971). Pupillary responses during a short-term memory task: Cognitive processing, arousal or both? *Journal of Experimental Psychology*, **90**, 311-318.
- KAHNEMAN, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.
- KELLAS, G., MCCAULEY, C., & MCFARLAND, C. E. (1975a). Developmental aspects of storage and retrieval. *Journal of Experimental Child Psychology*, **19**, 511-521.
- KELLAS, G., MCCAULEY, C., & MCFARLAND, C. E. (1975b). Reexamination of externalized rehearsal. *Journal of Experimental Psychology: Human Learning & Memory*, **1**, 84-90.
- KEPPEL, G. (1982). *Data analysis for research designs*. New York: W. H. Freeman.
- KUCERA, H., & FRANCIS, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- LOCKHART, R., & CRAIK, F. I. M. (1990). Levels of processing: A retrospective commentary on a framework for memory research. *Canadian Journal of Psychology*, **44**, 87-112.
- LOGAN, G. D. (1994). On the ability to inhibit thought and action: A users' guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language* (pp. 189-239). San Diego: Academic Press.
- LUPIÁÑEZ, J., & MILLIKEN, B. (1999). Inhibition of return and the attentional set for integrating versus differentiating information. *Journal of General Psychology*, **126**, 392-418.
- MACLEOD, C. M. (1989). Directed forgetting affects both direct and indirect tests of memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 13-21.
- MACLEOD, C. M. (1998). Directed forgetting. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 1-57). Mahwah, NJ: Erlbaum.
- MACLEOD, C. M. (1999). The item and list methods of directed forgetting: Test differences and the role of demand characteristics. *Psychonomic Bulletin & Review*, **6**, 123-129.
- 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). San Diego: Academic Press.
- MARKS, W., & DULANEY, C. L. (2001). Encoding processes and attentional inhibition in directed forgetting. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 1464-1473.
- MARTIN, D. W., & KELLY, R. T. (1974). Secondary task performance during directed forgetting. *Journal of Experimental Psychology*, **103**, 1074-1079.
- MCNALLY, R. J., CLANCY, S. A., BARRETT, H. M., & PARKER, H. A. (2004). Inhibiting retrieval of trauma cues in adults reporting histories of childhood sexual abuse. *Cognition & Emotion*, **18**, 479-493.
- MUTHER, W. S. (1965). Erasure of partitioning in short-term memory. *Psychonomic Science*, **3**, 429-430.
- OVERTOOM, C. C. E., KENEMANS, J. L., VERBATEN, M. N., KEMNER, C., VAN DER MOLEN, M. W., VAN ENGELAND, H., ET AL. (2002). Inhibition in children with attention-deficit/hyperactivity disorder: A psychophysiological study of the stop task. *Biological Psychiatry*, **51**, 668-676.
- PALLER, K. A. (1990). Recall and stem-completion priming have different electrophysiological correlates and are modified differentially by directed forgetting. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **16**, 1021-1032.
- PLISZKA, S. R., LIOTTI, M., & WOLDORFF, M. G. (2000). Inhibitory control in children with attention-deficit/hyperactivity disorder: Event-related potentials identify the processing component and timing of an impaired right-frontal response-inhibition mechanism. *Biological Psychiatry*, **48**, 238-246.
- REBER, P. J., SIWIEC, R. M., GITLEMAN, D. R., PARRISH, T. B., MESULAM, M. M., & PALLER, K. A. (2002). Neural correlates of successful encoding identified using functional magnetic resonance imaging. *Journal of Neuroscience*, **22**, 9541-9548.
- RIEGER, M., GAUGGEL, S., & BURMEISTER, K. (2003). Inhibition of ongoing responses following frontal, nonfrontal, and basal ganglia lesions. *Neuropsychology*, **17**, 272-282.
- ROEDIGER, H. L., III, & CROWDER, R. G. (1972). Instructed forgetting: Rehearsal control or retrieval inhibition (repression)? *Cognitive Psychology*, **3**, 244-254.
- RUNDUS, D., & ATKINSON, R. C. (1970). Rehearsal processes in free recall: A procedure for direct observation. *Journal of Verbal Learning & Verbal Behavior*, **9**, 99-105.
- SAHAKYAN, L., & DELANEY, P. (2005). Directed forgetting in incidental learning and recognition testing: Support for a two-factor account. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **31**, 789-801.
- SAHAKYAN, L., & KELLEY, C. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 1064-1072.
- SAMUEL, A. G., & KAT, D. (2003). Inhibition of return: A graphical meta-analysis of its time course and an empirical test of its temporal and spatial properties. *Psychonomic Bulletin & Review*, **10**, 897-906.
- SEGO, S. A., GOLDING, J. M., & GOTTLÖB, L. R. (2006). Directed forgetting in older adults using the item and list methods. *Ageing, Neuropsychology, & Cognition*, **13**, 95-114.
- SHEARD, E. D., & MACLEOD, C. M. (2005). List method directed forgetting: Return of the selective rehearsal account. In N. Ohta, C. M. MacLeod, & B. Uttl (Eds.), *Dynamic cognitive processes* (pp. 219-248). Tokyo: Springer.
- TAYLOR, T. L. (2005). Inhibition of return following instructions to remember and forget. *Quarterly Journal of Experimental Psychology*, **58A**, 613-629.
- TAYLOR, T. L., & FAWCETT, J. M. (2008). *Larger IOR effects following forget than following remember instructions depend on endogenous attentional withdrawal and target localization*. Manuscript submitted for publication.
- TAYLOR, T. L., & KLEIN, R. M. (2000). Visual and motor effects in IOR. *Journal of Experimental Psychology: Human Perception & Performance*, **26**, 1639-1656.
- TEKAN, A. I., & AKTÜRK, M. (2001). Are you sure you forgot? Feeling of knowing in directed forgetting. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 1487-1490.
- ULLSPERGER, M., MECKLINGER, A., & MÜLLER, U. (2000). An electrophysiological test of directed forgetting: The role of retrieval inhibition. *Journal of Cognitive Neuroscience*, **12**, 924-940.
- WOODWARD, A., BJORK, R., & JONGEWARD, R. (1973). Recall and recognition as a function of primary rehearsal. *Journal of Verbal Learning & Verbal Behavior*, **12**, 608-617.
- WYLIE, G. R., FOXE, J. J., & TAYLOR, T. L. (2008). Gone but not forgotten: An fMRI investigation into the fate of discarded information in a directed forgetting paradigm. *Cerebral Cortex*, **8**, 670-682.
- ZACKS, R. T., & HASHER, L. (1994). Directed ignoring: Inhibitory regulation of working memory. In D. Dagenbach & T. H. Carr (Eds.), *Inhibitory processes in attention, memory, and language*. San Diego: Academic Press.
- ZACKS, R. T., RADVANSKY, G., & HASHER, L. (1996). Studies of directed forgetting in older adults. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 143-156.

## NOTES

1. Interestingly, if cumulative rehearsal is responsible for the equation of post-R and post-F probe RTs at the longest SOA, there is no indication that it becomes more cognitively demanding as the trials progress. We divided the study block into thirds and analyzed the postinstruction probe RTs as a function of memory instruction (R, F), SOA (1,400, 1,800, 2,600 msec), and block segment (Trials 1-40, Trials 41-80, Trials 81-120). There was no main effect of block segment ( $F < 1$ ) and no interaction of block segment with either memory instruction or SOA (both  $ps > .10$ ). Perhaps this is not surprising in light of the fact that each block segment included an average of approximately 20 R words. It is unlikely that the intertrial interval was long enough to allow participants to rehearse *all* R words cumulatively, regardless of whether they were in the first or last block segment. As such, a subset of R words was likely selected on each trial, with rehearsal of this subset being equally demanding at all except possibly several of the very earliest trials.

2. Note that because postword probe responses were not analyzed in Experiment 1, any interference of the memory instruction on the 2,600-msec postword probe RT was irrelevant to our analysis. Potential

overlap between the last probe presentation prior to instruction onset was corrected in Experiment 2; however, this was done under the assumption that the postword (or X-string) probe RTs might provide a suitable baseline for eliminating the effects of cumulative rehearsal from the postinstruction probe RTs.

3. Of the 38 participants, 4 were tested twice accidentally for a small number of words (around 12) during the recognition phase. Importantly, these duplicated recognition trials were evenly distributed across R and F conditions. For analyses, the response made to the second instance of each duplicate presentation was not scored; however, whether such du-

plicated trials were ignored or included had no effect on the significance or magnitude of the directed forgetting effect (both R and F proportions rounded to the same values regardless of whether these duplicated trials were included). These duplications occurred only during the recognition phase and did not affect any of the study phase trials (or, therefore, probe RTs).

(Manuscript received May 19, 2007;  
revision accepted for publication April 20, 2008.)