Acta Psychologica 144 (2013) 332–343

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

Event-method directed forgetting: Forgetting a video segment is more effortful than remembering it

Jonathan M. Fawcett [a,*], Tracy L. Taylor [a], Lynn Nadel [b]

[a] Dalhousie University, Halifax, Nova Scotia, Canada
[b] University of Arizona, Tucson, AZ, United States

ARTICLE INFO

Article history:
Received 22 September 2012
Received in revised form 16 April 2013
Accepted 7 July 2013
Available online 7 August 2013

PsycINFO classification:
2300 Human Experimental Psychology
2340 Cognitive Processes
2343 Learning & Memory

Keywords:
Intentional forgetting
Memory
Cognition
Events

ABSTRACT

Videos were presented depicting events such as baking cookies or cleaning a fish tank. Periodically, the video paused and an instruction to Remember (R) or Forget (F) the preceding video segment was presented; the video then resumed. Participants later responded more accurately to cued-recall questions (E1) and to true/false statements (E2–5) regarding R segments than F segments. This difference was larger for specific information (*the woman added 3 cups of flour*) than for general information (*the woman added flour*). Participants were also slower to detect visual probes presented following F instructions compared to those presented following R instructions. These findings suggest that intentional forgetting is an effortful process that can be performed even on segments of otherwise continuous events and that the result is a relatively impoverished representation of the unwanted information in memory.

1. Introduction

In a typical directed-forgetting paradigm, participants are presented with a series of study items (words, pictures, sentences, etc.) each of which they are instructed to either remember (R) or forget (F). When memory is subsequently tested for both the R and F items, participants perform better for R items compared to F items. This difference is referred to as a directed forgetting effect. There have been many variants of the directed forgetting paradigm that differ primarily in terms of how and when the R and F instructions are presented (for a review see Basden & Basden, 1998 or Bjork, 1972). Most of these variants have been categorized as belonging to either the item method or the list method (see Basden, Basden, & Gargano, 1993; for a review, see MacLeod, 1998).

The item-method paradigm, first developed by Bjork and Woodward (1973), is generally employed to study intentional forgetting at encoding. In this paradigm, study items are presented one at a time, each followed by an R or F instruction (e.g., Hourihan, Ozubko, & MacLeod, 2009; MacLeod, 1989; Quinlan, Taylor, & Fawcett, 2010); memory is subsequently tested for all items. A directed forgetting effect obtained using the item method is attributed to selective rehearsal of the R over the F items (Basden et al., 1993), accomplished in part by the engagement of cognitive mechanisms that actively withdraw processing resources from the representation of the unwanted F item in working memory (Fawcett & Taylor, 2010; Taylor, 2005; Taylor & Fawcett, 2011) and any other items that enter working memory shortly thereafter (Fawcett & Taylor, 2012). In contrast, the list-method paradigm is generally employed to study intentional forgetting at retrieval. In this paradigm, a single R or F instruction is presented following study of a complete list of items, after which participants are asked to remember a second list (e.g., Geiselman, Bjork, & Fishman, 1983; McNally, Clancy, Barrett, & Parker, 2004); a directed forgetting effect is measured as better subsequent memory for List 1 items for participants receiving an R instruction rather than an F instruction (referred to as the *costs* of directed forgetting) as well as superior performance for List 2 items for participants receiving an F instruction rather than an R instruction (referred to as the *benefits* of directed forgetting). Although the directed forgetting effect obtained using the list method has historically been attributed to inhibition of the F list at retrieval (e.g., Basden et al., 1993; Geiselman et al., 1983), more recent data suggests that a change in mental context between the R and F list presentation may play at least some role (Sahakyan & Kelley, 2002). Importantly, inhibitory and context-driven accounts of list-method directed forgetting need not be mutually exclusive and many modern theorists still believe inhibition to play a crucial role in this paradigm (e.g., Pastotter, Kliegl, & Bauml, 2012; Racsmány & Conway, 2006).

Recognizing the limitations of these traditional paradigms, Golding and Keenan (1985) and later Gottlob, Golding, and Hauselt (2006)

* Corresponding author at: Dalhousie University, Department of Psychology, Halifax, NS B3H 4J1, Canada. Tel.: +1 902 494 3001.
E-mail address: jmfawcet@dal.ca (J.M. Fawcett).

0001-6918/$ – see front matter
http://dx.doi.org/10.1016/j.actpsy.2013.07.005

explored whether participants could forget part of a continuous narrative. Golding and Keenan (1985) found that participants remembered erroneous spatial directions to ensure that these errors were not incorporated into future navigational decisions. In other words, marking the directions as irrelevant did not lead to intentional forgetting, as happens when an F instruction is applied to discrete words or lists of words, purportedly because whereas that information became nominally irrelevant to the navigational task, being erroneous did not render the directions functionally irrelevant as they could still prevent a wrong turn. In contrast, Gottlob et al. (2006) found that participants were capable of intentionally forgetting phone numbers that had been labeled as erroneous and replaced by the "correct" number. In this case, there was no inherent value to remembering the irrelevant phone number so it was successfully forgotten.

Joslyn and Oakes (2005) conducted the first diary study of directed forgetting in which participants kept a written record of the events they experienced across a two-week period. After the first week, half of the participants were instructed to forget the entries they had recorded whereas the remaining participants were given no such instruction. The participants who received an F instruction recalled the descriptive titles of fewer Week 1 events compared to the participants who did not receive this instruction. No difference was observed for the details of the events for which participants successfully retrieved the title, although it is possible that retrieval of the title may have resulted in a release of inhibition (e.g., Bjork & Bjork, 1996) and/or reinstatement of context (e.g., Sahakyan & Kelley, 2002) masking any differences for this measure. While innovative, it is possible that participants recorded central details pertaining to the reported events making them more readily retrieved so long as the titles of the events were available. This issue is not easily resolved using self-generated information but rather requires the presentation of events under controlled conditions.

Fawcett, Taylor, and Nadel (in press) addressed this concern using a novel *event-method* directed forgetting paradigm in which they embedded R and F instructions into videotaped vignettes that depicted a continuous sequence of events aimed at accomplishing a single goal (e.g., baking cookies).[1] In their study participants watched four videos depicting common events (e.g., such as baking cookies) during which they were instructed to remember certain segments of the otherwise continuous event and forget others. Each video consisted of eight segments lasting 35 s that were presented sequentially without interruption so that, from the participants' perspective, the video was a continuous sequence of events. Memory instructions were represented by changing the color of the border that surrounded the viewing port containing the video: Participants were required to remember everything that was presented in the video while the border was green and to forget everything that was presented in the video while the border was purple. The assignment of the R and F instructions was randomized across segment, with the restriction that each video contained four R segments and four F segments.

Across five experiments, Fawcett et al. (in press) observed better subsequent memory performance for R segments compared to F segments using test questions or true/false statements. This difference remained even when an event segmentation task (see Zacks, Tversky, & Iyer, 2001) was employed to encourage conceptual encoding of the *entire* video (i.e., all R and F segments). In their final experiment, Fawcett et al. (in press) demonstrated that the effect of intentional forgetting in this paradigm was smaller (or even non-existent) for relatively general test statements (e.g., *the woman added flour*) compared to the robust effect observed for relatively specific test statements (e.g., *the woman added 3 cups of flour*). This finding suggests that intentional forgetting has a graded effect on the to-be-forgotten information, with a greater loss of details relative to gist (although see Joslyn & Oakes, 2005).

Fawcett et al. (in press) provided a strong test of the hypothesis that participants could selectively forget the details of unwanted events when the memory instructions were presented concurrent to the studied material. Concurrent memory instructions unobtrusively indicated the R and F information without interrupting the events to which they referred and therefore emulated a natural viewing experience. However, this finding would be ever more compelling if demonstrated in a paradigm wherein the memory instruction was presented after the to-be-remembered or to-be-forgotten segment had already been encoded. Whereas a concurrent memory instruction requires the participant to control the manner in which the R or F information is encoded, a delayed memory instruction requires the participant to control the representation of the R or F information within memory. Accordingly, Fawcett et al. (in press) demonstrated that participants could preferentially ignore F segments and process R segments as they were encoded, impacting the specificity of the resulting memory trace. It is our current goal to determine whether participants are capable of preferentially suppressing F segments and processing R segments immediately after they have been encoded — and whether this effect will also be limited to relatively specific information. To address this question the current experiment adapted Fawcett et al.'s (in press) paradigm to use a delayed as opposed to a concurrent memory instruction: Following each segment, the video paused, the screen cleared and participants received a green- or purple-filled circle instructing them to remember or forget the *preceding* segment. Further, to explore the mechanisms via which the R and F instructions are instantiated in our task we presented a visual probe ('*') requiring a speeded detection response following most of the R and F memory instructions (see Fawcett & Taylor, 2008; see also, Fawcett & Taylor, 2010, 2012; Taylor, 2005; Taylor & Fawcett, 2011).

2. Experiment 1

In Experiment 1, participants viewed videos of common events such as baking cookies or cleaning a fish tank. The videos were each separated into eight discrete segments lasting 35 s: Participants were instructed to remember a random half of the segments contained within each video and to forget the remainder. Participants were also required to make a speeded response to report the detection of a probe sometimes presented following the memory instruction. Longer reaction times (RTs) were taken as an index of increasing cognitive demands (see Kahneman, 1973). Following the study phase trials, participants responded to questions testing their knowledge for *all* video segments regardless of the associated memory instruction. Recent evidence within the item-method has revealed that enacting an F instruction is an effortful process capable of slowing subsequent responses (e.g., Fawcett & Taylor, 2008), interacting with visual attention through the magnification of inhibition of return (Taylor, 2005; Taylor & Fawcett, 2011), and interfering with the formation of incidental memories (Fawcett & Taylor, 2012). These behavioral findings, along with recent neuroimaging work (e.g., Wylie, Foxe, & Taylor, 2008) suggest that intentional forgetting may under certain circumstances involve the engagement of active control processes aimed at suppressing further processing of the unwanted information. In light of these findings, in addition to predicting that participants would respond more accurately when tested for R compared to F segments, to the extent that instantiating an F instruction is more effortful than instantiating an R instruction (Fawcett & Taylor, 2008) we also

[1] We have adopted the term *event-method directed forgetting* to describe Fawcett et al. (in press) as well as the current paradigm because it emphasizes the target of the R and F memory instructions. Whereas item-method directed forgetting pairs each memory instruction with a specific item and list-method directed forgetting pairs each memory instruction with a specific list, the memory instructions in the current study cannot be ascribed to either. Each segment is no more an item than a list. The segments instead represent the dynamic combination of visual features into a cohesive vignette with numerous sub-elements that are broadly conceptualized as "events". In adopting this terminology we recognize that certain past experiments also fall within this definition (e.g., Joslyn & Oakes, 2005) — we do not claim to be the first to study the intentional forgetting of events or actions, we only intend to encourage others to recognize that the item/list-method nomenclature is perhaps unbefitting to such instances.

predicted that they would exhibit slower probe RTs following F than R instructions.

3. Method

3.1. Participants

Twenty-nine undergraduate students (18 female) enrolled at Dalhousie University or the University of Arizona participated in this experiment for course credit. Most participants were right-handed (18 right).

3.2. Stimuli and apparatus

The experimental task used a custom script developed in the Python programming language (www.python.org) with the Pygame development library (www.pygame.org). The script was loaded on either a 17″ MacBook Pro computer running Mac OS X 10.5 or a 24″ iMac computer running Mac OS X 10.5. Responses were recorded via the built-in laptop keyboard. Instructions and test statements were presented against a black background in white, size 18 Gentium Basic Bold (www.sil.org/~gaultney/Gentium/); the title of each video was instead presented in size 30 of the same font.

Four videos depicting common events (Baking Assorted Cookies, Getting Ready for Work, Cleaning a Fish Tank, Making Chocolate Pudding) were retrieved from the public domain video sharing website YouTube and used during the study phase; a 5th video (Doing Laundry) was retrieved and used during a preceding practice phase. Each video contained a linear progression of events resulting in a predetermined, self-evident goal as described by the title of that video (e.g., "Baking Assorted Cookies"). Videos were resized to fit into an area measuring 600×600 pixels and were composed of 7000 frames (1750 frames for the practice video). Videos were presented at an average rate of 25 frames per second, for a total duration of 4 min and 40 s (1 min and 10 s for the practice video). Throughout the practice and study phase, participants were instructed to remember (R) a random half of segments from each video and to forget the remaining (F) segments in accordance with green- or purple-filled circles presented following each segment. Each circle measured 75 pixels in diameter and was surrounded by a 3-point white border. Each segment lasted 875 frames (35 s), resulting in 8 segments per video (2 segments for the practice video). R and F instructions were randomly assigned to the ordered sequence of video segments on a subject-by-subject basis with the restriction that each video always contained 4 R segments and 4 F segments (1 R segment and 1 F segment for the practice video).

Test questions were used to probe memory for the contents of each video. Each question tested information revealed only during a single segment. Because of inherent differences in the amount of unique information contained within the segments, two of the videos (Baking Assorted Cookies, Making Chocolate Pudding) were tested with 3 test questions per segment and the remaining two videos (Cleaning a Fish Tank, Getting Ready for Work) were tested with 2 test questions per segment for a total of 80 questions overall (see Fawcett et al., in press, Experiment 1).

3.3. Procedure

Participants were told that they would view a series of videos each depicting a common event – such as baking cookies – some portion of which they would be instructed to remember for a subsequent memory test.

3.3.1. Practice phase

A practice video about folding laundry was presented to familiarize participants with the task. During this video the experimenter provided sample questions so that the participant would understand the type of information they were expected to retain. The presentation of the practice video was identical to the study phase trials except that only a single R segment and a single F segment were presented. Participants were then presented with a written version of the study phase instructions on the computer screen and told to press 'RETURN' when they were ready to begin the experiment proper.

3.3.2. Study phase

Prior to each video a descriptive title (e.g., "Baking Assorted Cookies") was presented in the center of the screen until the participant pressed the 'RETURN' key at which point the video began. Videos were separated into 8 segments each lasting 35 s and followed by a green- or purple-filled circle instructing the participant to remember or forget that segment: Each video contained a total of 4 R and 4 F segments. At the end of each 35 s segment, the video paused and the screen cleared. Following a 3000 ms delay, the memory instruction was presented for 300 ms and then removed resulting in a blank screen. To assess relative cognitive load associated with instantiating the R and F instructions, on 75% of trials, a single visual probe ("*") lasting 400 ms was presented at a stimulus onset asynchrony (SOA) of 1800 ms in relation to onset of the memory instruction; participants made a speeded detection response to this probe by pressing the spacebar as quickly as possible. We selected 1800 ms for the instruction-probe SOA because Fawcett and Taylor (2008) found the F > R differences in probe RTs to be most robust at this interval. On the remaining 25% of the trials (no-probe catch trials) no visual probe was presented. Catch trials were included to measure the false alarm rate associated with the detection response. Probe and catch trials were equally distributed across memory instruction and video.

Following the disappearance of the memory instruction there was a delay of 8700 ms before the video resumed, such that the total duration of the pause between sequential video segments was 12,000 ms. The timing of events within a single video is depicted in Fig. 1 and was repeated until all 8 segments (and their associated memory instructions) had been presented. The presentation order of the four study phase videos was randomized; the segments contained within each video were presented in sequential order, interrupted only by the 12,000 ms pause described above. Once a given video ended, the title for the next video was presented until the participant initiated play by pressing the 'RETURN' key. This process continued until all 4 videos were completed at which point the instructions for the test phase were displayed.

3.3.3. Test phase

During the test phase, participants were presented with a series of questions, individually in the center of the computer screen, each pertaining to a specific R or F segment. Participants were tested for the content of one video at a time; the order in which the videos were tested was randomized and the statements for a given video were presented in a random order. Participants were instructed to answer each question to the best of their ability using one or two words. To avoid participants dwelling on a given question, they were instructed to respond *dnk* (*do not know*) or *idk* (*I don't know*) if unable to even guess — although the use of this response was discouraged.

The first author scored each question using an answer key created prior to data collection. During scoring, the R or F instruction associated with the scored question was obfuscated to prevent bias. Responses that the first author felt were correct but were not listed on the original answer key were flagged for independent review by an assistant and either added to the answer key or rejected. Once all data had been collected, each question was rescored using *only* the answer key to ensure consistency. Misspellings were accepted as correct only if they were unambiguous. For example, *fingur* instead of *finger* would be acceptable whereas *air* instead of *hair* would not be acceptable.

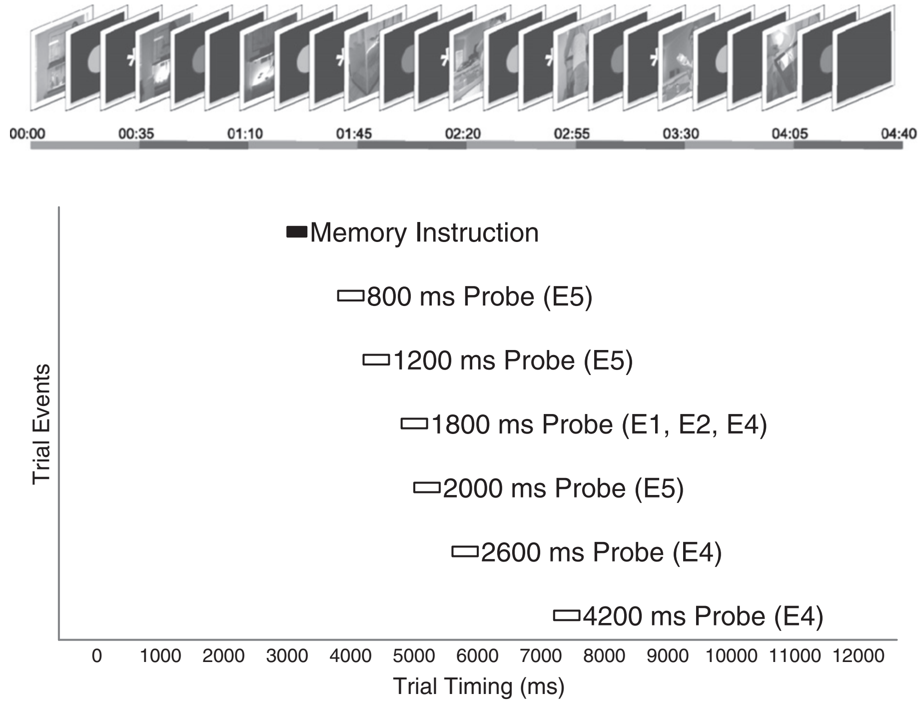

Fig. 1. The top panel depicts a schematic representation of the study phase presentation of "Cleaning a Fish Tank" with a complementary timeline denoting the start and the end time of each segment if the video was to run uninterrupted. Each video segment corresponds to one bar within the timeline; in the figure, a single still frame capture is used to represent the 35 s segment. For the sake of this depiction, the light circles represent R instructions and the dark circles represent F instructions. R and F instructions were randomly assigned to each segment on a subject-by-subject basis such that each video contained four R segments and four F segments. For the purpose of this example only, the R and F segments are shown as alternating one after the other. The bottom panel depicts the trial events that occurred during the pause following each segment and timed in relation to video offset (which is denoted as 0 ms). Only a single probe was presented on each study phase trial with the exception of no-probe catch trials. Experiment 3 did not contain any probes.

4. Results

4.1. Recall data

The percentage of correct responses made on the test was calculated by dividing the number of responses matching the answer key by the total number of questions answered; *do not know* responses were treated as incorrect, since the participant had asserted that they did not know the correct answer — excluding these responses did not change the results. The percentage of correct responses was then analyzed as a function of instruction (R, F) using a one-way repeated-measures ANOVA. This analysis was significant, $F(1, 29) = 8.17$, $MSe = 39.81$, $p < .01$, $\eta^2_g = .065$, with better performance for test questions pertaining to R segments ($M = 49.11\%$, $SE = 1.78\%$) than F segments ($M = 44.38\%$, $SE = 1.60\%$). The magnitude of this directed-forgetting effect, 4.73%, was not significantly different from the 5.60% directed forgetting observed by Fawcett et al. (in press) using concurrent as opposed to delayed memory instructions, $t(57) = 0.30$, $p > .77$.

4.2. Probe detection RTs

A response on a probe-present detection trial was considered correct if the participant executed a response within 100 ms and 2000 ms of probe onset; correct detection responses were made on 96.84% of F trials ($SE = 1.62\%$) and on 99.14% ($SE = 0.48\%$) of R trials, which did not differ significantly $F(1, 28) = 2.22$, $MSe = 34.47$, $p = .147$, $\eta^2_g = .032$. Mean probe RTs for correct trials were analyzed as a function of instruction (R, F) using a repeated-measures ANOVA. This analysis revealed probe RTs to be significantly longer following F instructions ($M = 532$ ms, $SE = 26$ ms) than R instructions ($M = 475$ ms, $SE = 24$ ms), $F(1, 28) = 18.35$, $MSe = 2595.96$, $p < .01$, $\eta^2_g = .044$. This finding replicates the F > R probe RT difference observed by Fawcett and Taylor (2008) in an item-method paradigm.

4.3. Probe false alarms

Finally, we analyzed the percentage of false alarms made on no-probe catch trials as a function of instruction (R, F). There were numerically fewer false alarms committed for F trials ($M = 2.59\%$, $SE = 1.44\%$) compared to R trials ($M = 3.44\%$, $SE = 1.63\%$) (see also Fawcett & Taylor, 2008); however, this difference was not significant, $F(1, 28) = 0.32$, $MSe = 33.10$, $p = .57$, $\eta^2_g = .003$.

5. Discussion

Despite presenting the memory instruction after the studied material, a significant directed forgetting effect was observed as measured by test questions. Further, the finding that participants were slower to respond to probes presented following an F than following an R instruction suggests that intentional forgetting in this paradigm may require engagement of an effortful cognitive process that helps limit further processing of the F segment. Fawcett and Taylor (2008) identified a similar pattern of F > R probe RTs for up to 1800 ms after instruction onset in an item-method paradigm. They argued that enacting an F instruction was associated with a brief, effortful process that discouraged further processing of the to-be-forgotten information. This effortful process may involve activation of frontal control mechanisms (see Wylie et al., 2008) that limit further rehearsal of F items (Fawcett & Taylor, 2008) – as well as items that appear shortly thereafter (Fawcett & Taylor, 2012) – likely by withdrawing attentional resources from the F item representation during encoding (Taylor, 2005; see also, Fawcett & Taylor, 2010; Taylor & Fawcett, 2011). The current findings provide the first evidence that similar mechanisms may be engaged flexibly to prevent the encoding of segments of otherwise continuous events.

6. Experiment 2

Having demonstrated a directed forgetting effect using recall we next explored whether memory instruction affects the quality of the event after encoding. To address this question, Experiment 2 replicated the methods of Experiment 1 with the exception that memory was tested using relatively specific or general true/false statements. Whereas directed forgetting is typically measured as an all-or-nothing phenomenon based on a comparison of overall memory performance for R and F information, by manipulating the relative specificity of the test statements, we explored whether intentional forgetting differentially impacts the details and the gist of the targeted event.

7. Methods

7.1. Participants

Thirty undergraduate students (19 female) enrolled at Dalhousie University or the University of Arizona participated in this experiment for course credit. Most participants were right-handed (25 right).

7.2. Stimuli and apparatus

The stimuli and apparatus used in the current experiment were identical to those used in Experiment 1 with the exception that true/false statements were presented instead of test questions. Eight true/false test statements were created for each segment. True statements were created first and referred to a particular event (e.g., *the pudding was served in a clear glass with a stem*) or fact (e.g., *the recipe called for 2 tablespoons of cornstarch*) revealed only during the relevant segment. False statements were most often created by replacing a single detail within each true statement, maintaining the general structure whenever possible (e.g., *the recipe called for 2 tablespoons of salt*); in some cases, this was not possible. General statements sometimes were created by removing details from the specific statements (e.g., *the recipe called for cornstarch*) although in other cases alternate information was tested. Overall, 256 test statements (128 true and 128 false) were created and were equally distributed across specificity (128 specific and 128 general).

7.3. Procedure

The procedure was identical to Experiment 1 with the exception that instead of questions at test, participants were presented with a series of true/false statements. Participants were tested for the content of one video at a time, although the statements for that video were presented in a random order. The specific and general statements as well as the true and false statements were randomly interspersed. The title of the video being tested was presented directly above each test statement. Participants pressed “j” on the computer keyboard to indicate a statement that was true or “f” to indicate that the statement was false (with the mnemonic that “f” was for false to ensure participants did not confuse this response mapping). Responses were self-paced and no feedback was given.

8. Results

8.1. Signal detection analysis

The raw hits and false alarms are provided in Table 1. Using the procedure described by Donaldson (1992), non-parametric measures of sensitivity (A') and response bias (B''_D) were calculated and analyzed as a function of instruction (R, F) and relative specificity (specific, general) using separate two-way repeated-measures ANOVAs. These data are provided in Fig. 2. For the A' analysis, the main effect of instruction was significant with greater sensitivity to statements about R segments than F segments, $F(1, 29) = 4.22$, $MSe = 0.005$, $p = .049$, $\eta^2_g = .028$. The main effect of specificity was also significant with greater sensitivity to general statements than specific statements, $F(1, 29) = 183.22$, $MSe = 0.004$, $p < .001$, $\eta^2_g = .536$. Importantly, these effects were qualified by a significant instruction × specificity interaction, $F(1, 29) = 7.50$, $MSe = 0.003$, $p = .010$, $\eta^2_g = .033$. Planned contrasts revealed a significant 0.05 directed forgetting effect in A' for specific test statements, $t(29) = 2.76$, $p = .010$. Not only was there no directed forgetting effect in A' for general test statements, but also there was a non-significant 0.01 *reverse* directed forgetting effect in this condition, $t(29) = 0.18$, $p = .862$.

The B''_D analysis revealed only a significant main effect of relative specificity, $F(1,29) = 1.26$, $MSe = 0.06$, $p < .001$, $\eta^2_g = .156$, with participants employing a more liberal response bias for general statements than specific statements; neither the main effect of instruction, $F(1,29) = 1.26$, $MSe = 0.08$, $p = .270$, $\eta^2_g = .008$, nor the instruction × specificity interaction, $F(1,29) = 0.19$, $MSe = 0.05$, $p = .665$, $\eta^2_g < .001$, was significant.

8.2. Probe detection RTs

A probe-present detection trial was considered correct if the participant executed a response within 100 ms and 2000 ms of probe onset; correct detection responses were made on 97.78% of F trials ($SE = 1.19\%$) and on 99.72% ($SE = 0.28\%$) of R trials, which did not differ significantly from one another, $F(1, 29) = 2.45$, $MSe = 23.19$, $p = .129$, $\eta^2_g = .042$. Mean probe RTs for correct trials were analyzed as a function of instruction (R, F) using a repeated-measures ANOVA. This analysis revealed probe RTs to be significantly longer following F instructions ($M = 504$ ms, $SE = 34$ ms) than R instructions ($M = 460$ ms, $SE = 32$ ms), $F(1, 29) = 10.52$, $MSe = 2784.18$, $p = .003$, $\eta^2_g = .015$.

8.3. Probe false alarms

The percentage of false alarms was numerically equivalent for R trials ($M = 2.50\%$, $SE = 1.39\%$) and for F trials ($M = 2.50\%$, $SE = 1.84\%$) and so were not significantly different, $F < 1$.

9. Discussion

Experiment 2 replicated Experiment 1 with the exception that true/false statements were used instead of test questions. Our signal detection analysis revealed a significant directed forgetting effect for

Table 1
Mean percentage of “true” responses (i.e., hits for true statements and false alarms for false statements) as a function of instruction (remember, forget), specificity (specific, general) and statement validity (true, false) in Experiments 2, 3, 4 and 5. Control denotes performance for the relevant remember-all group.

Instruction	Specific				General			
	True		False		True		False	
	M	SE	M	SE	M	SE	M	SE
Experiment 2								
R	67.92	1.55	28.44	2.10	88.13	1.40	17.29	1.73
F	61.67	2.21	31.56	2.07	86.88	2.12	17.29	1.72
Control	63.40	1.57	28.13	1.25	89.30	0.94	17.12	1.39
Experiment 3								
R	67.66	1.86	32.50	2.41	87.03	1.62	23.44	2.74
F	63.59	2.18	37.50	2.79	88.75	1.53	24.69	2.88
Control	65.51	2.69	32.03	3.33	86.83	1.34	22.32	3.43
Experiment 4								
R	63.15	1.87	29.04	1.87	88.67	1.39	18.49	2.32
F	59.77	1.66	32.55	1.67	88.54	1.44	18.62	1.52
Experiment 5								
R	70.31	1.77	28.82	1.83	83.68	2.46	19.62	1.75
F	62.85	2.51	29.69	2.19	84.38	2.29	19.44	2.43

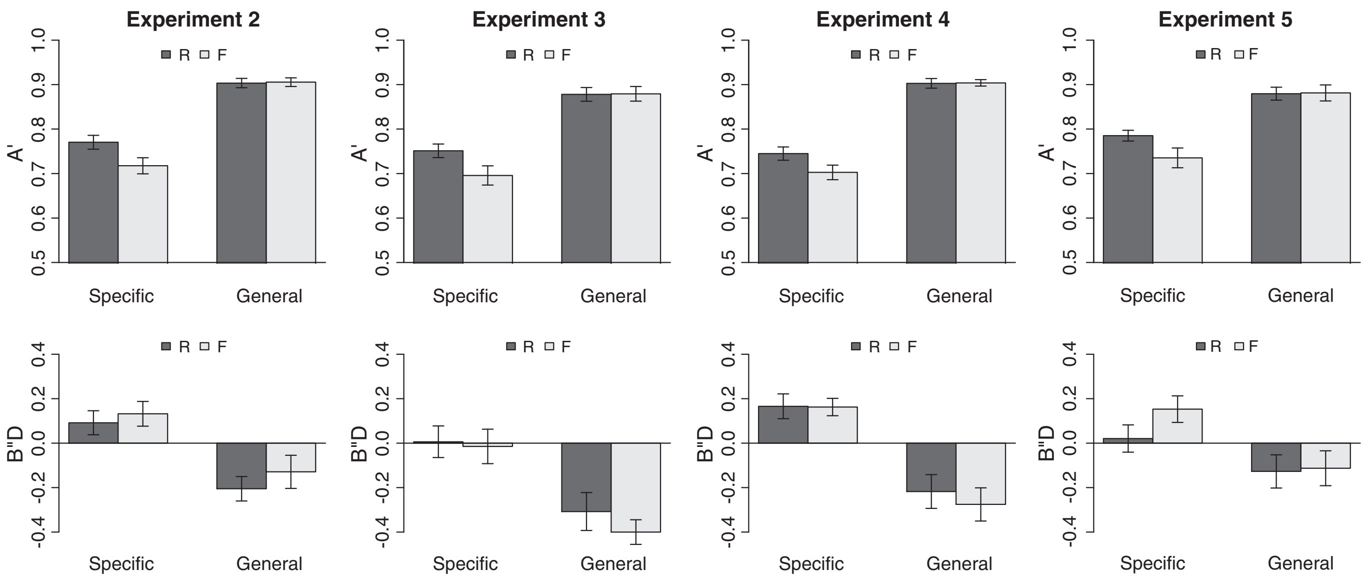

Fig. 2. Sensitivity (A') and Response bias (B''_D) as a function of instruction (R, F) and specificity (specific, general) for Experiments 2, 3, 4 and 5; error bars represent one standard error of the mean.

specific test statements as measured by A'. Instruction did not affect sensitivity to general test statements and there was no effect of instruction on response bias (B''_D) for either level of specificity. The interaction between directed forgetting and relative specificity for A' suggests that participants are capable of using a memory instruction to control the level of representation that is encoded for recently viewed events: R items are represented with greater detail than F items, with no effect of instruction on the gist representation.

Because the general test statements necessarily queried information that might otherwise be part of a general schema activated by the video title (e.g., "Baking Assorted Cookies"), it is important to know whether the failure to observe a directed forgetting effect for general statements (e.g., *the woman added cornstarch*) was due to the fact that participants responded based on schema activation (e.g., making assorted cookies usually requires adding cornstarch) rather than based on memory for the presented video events (e.g., remembering that the woman in the video added cornstarch). Data from a control condition run by Fawcett et al. (in press) speak directly to this issue. Using the same test statements presented above, Fawcett et al. (in press) had 10 participants guess the correct answer for each, based only on the title of the event; in this "no-study" control condition, participants did not view the videos prior to answering the test statements. If participants had been able to guess the answer in many of the general test statements, this control group would have shown relatively high sensitivity (i.e., a high A' score) to general test statements despite not having watched the videos. In fact, performance in this no-study control condition was equivalent to chance for specific test statements ($M = 0.50$, $SE = 0.02$) and only slightly better than chance for the general test statements ($M = 0.59$, $SE = 0.04$). And, more to the point, in both cases, the A' score for Fawcett et al.'s (in press) no-study control group was drastically lower than the A' score obtained for the specific ($M_R = .77$ and $M_F = .72$) and general test items ($M_R = .90$ and $M_F = .91$) in the current experiment. This provides evidence that participants responded to the general test statements based on memory, and not schema activation, such that the failure to obtain a directed forgetting effect suggests equivalent representation of gist for R and F traces.

Another control condition sometimes discussed in the directed forgetting literature involves presenting participants with an exact replication of the experimental task with the exception that participants are instructed to remember *all* study items as opposed to a subset of items (e.g., Basden & Basden, 1996; Sahakyan & Foster, 2009). To tease apart the costs and benefits, performance in each condition of the experimental task is compared to performance in the remember-all baseline task. Benefits are measured as better performance for R items in the experimental task compared to the remember-all baseline whereas costs are measured as worse performance for F items in the experimental task compared to the remember-all baseline (e.g., Basden & Basden, 1996; Sahakyan & Foster, 2009). Together the costs and benefits are thought to combine to produce what we have thus far referred to as the directed forgetting effect.

With this in mind, we collected data from an additional 26 participants, using methods exactly the same as in Experiment 2 with the exception that the colored circles were not ascribed any meaning; instead, participants were instructed to remember every segment of the videos they watched, while also responding to the visual probes. The raw hits and false alarms are provided in Table 1. This remember-all instruction resulted in mean A' scores of 0.76 ($SE = 0.01$) for the specific statements and 0.92 ($SE = 0.01$) for general statements. Planned t-tests comparing performance for the specific test statements between the baseline and experimental groups revealed marginally significant costs, $t(54) = 1.72$, $p = .090$, but no evidence of benefits, $t(54) = 0.76$, $p = .451$. For the general test statements performance was slightly higher in the remember-all baseline group than in either the R or F condition of the experimental group; this difference was not significant for either R segments, $t(54) = 0.97$, $p = .339$, or F segments, $t(54) = 0.84$, $p = .403$.

We also investigated the experimental probe RTs in relation to this remember-all condition. Participants in the remember-all task had an overall mean RT of 398 ms ($SE = 25$ ms). Although probe RTs were significantly shorter in the baseline task compared to F trials, $t(54) = 2.45$, $p = .018$, they did not differ significantly compared to R trials, $t(54) = 1.51$, $p = .136$. These findings converge with conclusions drawn from prior within-subject baseline comparisons conducted in the context of item-method directed forgetting. For example, in an item-method paradigm that incorporated reaction time probes, Fawcett and Taylor (2008, Experiment 2) interspersed trials on which a meaningless string of Xs was presented instead of a study word. This provided an online, within-subjects measurement of probe RTs on trials where there was no memory demand imposed by the presented item (i.e., the string of Xs). In a similar vein, Fawcett, Lawrence, and Taylor (2011) preceded the study phase of their experiment with a brief practice session with the probe task in which the same study item was repeated from trial-to-trial and no meaning was ascribed to the memory instruction. Responses to probes in this practice block provided a within-subjects measure of probe RTs in the absence of a memory demand. In both cases, probe RTs from the R trials roughly equated with the baseline probe RTs, whereas probe RTs from F trials tended to be longer.

The observation of slower probe RTs following F instructions compared to R instructions serves as a bridge between our event-method paradigm and the more classic item-method paradigm. Using the item method, Fawcett and Taylor (2008) also observed that within 1800 ms of memory instruction onset, probe RTs were slower following F than R instructions. They proposed that this pattern was an evidence of a brief cognitive mechanism that discouraged rehearsal of the discarded study item. Fawcett and Taylor (2012) provided further evidence that instantiating an item-method F instruction interferes with the formation of incidental memory for other, task-irrelevant items presented during the post-instruction period and Taylor (2005) (see also, Fawcett & Taylor, 2010; Taylor & Fawcett, 2011) observed a relatively larger inhibition of return effect for F trials relative to R trials. Evidence continues accumulating in support of an active mechanism of intentional forgetting that is associated with exerting control over the contents of working memory (see Fawcett & Taylor, 2012) with the possible benefit of subsequently biasing processing resources away from unreliable sources (see Taylor & Fawcett, 2011). The current findings provide preliminary evidence that the mechanism proposed in the context of intentionally forgetting discrete words or pictures is also relevant to continuous, contextually rich information such as provided by our video vignettes.

10. Experiment 3

Experiment 2 demonstrated a robust directed forgetting effect for specific but not general test statements. However, before we accept this finding we would like to ensure that it has not been influenced by the presence of the probe task. For example, it is conceivable that dedicating cognitive resources to the secondary probe detection task detracted from the resources available to rehearse the R segments or to forget the F segments. If so, the magnitude of the directed forgetting effect could be artificially reduced, masking any effects of the memory instruction on responses to the general test statements. To investigate this issue, Experiment 3 replicated Experiment 2 with the removal of the study phase probe response task.

11. Method

11.1. Participants

Twenty undergraduate students (14 female) enrolled at the University of Arizona or Dalhousie University participated in this experiment for course credit. Most participants were right-handed (18 right).

11.2. Stimuli and apparatus

The stimuli and apparatus were identical to Experiment 2.

11.3. Procedure

The procedure was identical to Experiment 2 with the exception that the probe response component of the experiment was removed from the study phase trials.

12. Results

The mean hits and false alarms are provided in Table 1. Non-parametric measures of sensitivity (A') and response bias (B''_D) were calculated and analyzed as a function of instruction (R, F) and relative specificity (specific, general) using separate two-way repeated-measures ANOVAs (see Fig. 2). For the A' analysis, there was a marginal effect of instruction with greater sensitivity to statements about R segments than F segments, $F(1, 19) = 3.76$, $MSe = 0.004$, $p = .068$, $\eta^2_g = .031$. The main effect of specificity was significant with greater sensitivity to general statements than specific statements, $F(1, 19) = 165.38$, $MSe = 0.003$, $p < .001$, $\eta^2_g = .511$. Importantly, these effects were qualified by a significant instruction × specificity interaction, $F(1, 19) = 7.47$, $MSe = 0.002$, $p = .013$, $\eta^2_g = .034$. Planned contrasts revealed a significant 0.06 directed forgetting effect in A' for specific test statements, $t(29) = 2.60$, $p = .018$. Once again, there was no directed forgetting effect in A' for general test statements, $t(29) = 0.09$, $p = .927$.

The B''_D analysis revealed only a significant main effect of relative specificity, $F(1, 19) = 38.37$, $MSe = 0.06$, $p < .001$, $\eta^2_g = .231$, with participants employing a more liberal response bias for general statements than specific statements; neither the main effect of instruction, $F(1, 19) = 1.10$, $MSe = 0.06$, $p = .307$, $\eta^2_g = .008$, nor the instruction × specificity interaction, $F(1, 19) = 0.59$, $MSe = 0.04$, $p = .452$, $\eta^2_g = .003$, was significant.

13. Discussion

Experiment 3 replicated the interaction between instruction and specificity for A' that was observed in Experiment 2, thus demonstrating that the presence of the probe task in Experiment 2 was not responsible for the interaction of the memory instruction and item specificity. As in Experiment 2, a remember-all baseline experiment was conducted with 14 new participants resulting in a mean A' value of 0.75 ($SE = 0.02$) for the specific statements and 0.89 ($SE = 0.02$) for the general statements (the raw hits and false alarms are provided in Table 1). Comparison of these values to performance in the experimental task revealed marginally significant costs, $t(32) = 1.79$, $p = .082$, and no benefits, $t(32) = 0.14$, $p = .892$, for specific test statements. Once again, performance for the general test statements was numerically equivalent for R and F segments so neither costs nor benefits were possible although performance was numerically (but not significantly) higher in the baseline group than in either the R condition, $t(32) = 0.31$, $p = .762$, or F condition, $t(32) = 0.25$, $p = .808$, of the experimental group.

Given that the pattern of memory performance did not differ between Experiments 2 and 3, a further analysis was conducted pooling these experiments as well as their baseline conditions to maximize statistical power. In this analysis the costs observed for the specific statements reached significance, $t(88) = 2.52$, $p = .013$, although the other comparisons remained non-significant as reported above (all $ps > .333$).

14. Experiment 4

Having clearly demonstrated that F segments are less specific in memory than R segments and that this is associated with an active mechanism following F instructions, Experiment 4 expanded upon the latter by exploring the time-course of the mechanism(s) indexed by the probe RTs. Fawcett and Taylor (2008) hypothesized that the processes associated with intentional forgetting resolved within approximately 2600 ms of instruction onset (for a replication, see Hansen, 2011). The few studies that have explored the time-course of the F > R probe RT difference stay within this temporal window (e.g., Fawcett & Taylor, 2012) so the time-course following 2600 ms remains unknown. Experiment 4 replicated Experiment 2 with the inclusion of 2600 ms and 4200 ms as well as 1800 ms instruction-probe SOAs to characterize the time-course at these later intervals.

15. Method

15.1. Participants

Twenty-four undergraduate students (15 female) enrolled at the University of Arizona or Dalhousie University participated in this experiment for course credit. Most participants were right-handed (22 right).

15.2. Stimuli and apparatus

The stimuli and apparatus were identical to Experiment 2.

15.3. Procedure

The procedure was identical to Experiment 2 with the exception that the visual probes were presented 1800 ms, 2600 ms or 4200 ms following memory instruction onset. Videos contained one of each SOA and one catch trial for each memory instruction. Presentation order was randomized on a subject-by-subject basis. The timing of the study phase trials otherwise remained unchanged.

16. Results

16.1. Signal detection analysis

The mean hits and false alarms are provided in Table 1. Non-parametric measures of sensitivity (A') and response bias (B''_D) were calculated and analyzed as a function of instruction (R, F) and relative specificity (specific, general) using separate two-way repeated-measures ANOVAs (see Table 1). For the A' analysis, there was a marginal effect of instruction with greater sensitivity to statements about R segments than F segments, $F(1, 23) = 4.12$, $MSe = 0.002$, $p = .054$, $\eta^2_g = .027$. The main effect of specificity was also significant with greater sensitivity to general statements than specific statements, $F(1, 23) = 292.32$, $MSe = 0.003$, $p < .001$, $\eta^2_g = .680$. Importantly, these effects were qualified by a significant instruction × specificity interaction, $F(1, 23) = 6.06$, $MSe = 0.002$, $p = .022$, $\eta^2_g = .030$. Planned contrasts revealed a significant 0.05 directed forgetting effect in A' for specific test statements, $t(23) = 2.66$, $p = .014$. There was no directed forgetting effect in A' for general test statements, $t(23) = 0.11$, $p = .910$.[2]

The B''_D analysis revealed only a significant main effect of relative specificity, $F(1, 23) = 40.96$, $MSe = 0.10$, $p < .001$, $\eta^2_g = .315$, with participants employing a more liberal response bias for general statements than specific statements; neither the main effect of instruction, $F(1, 23) = 0.28$, $MSe = 0.08$, $p = .601$, $\eta^2_g = .003$, nor the instruction × specificity interaction, $F(1, 23) = 0.43$, $MSe = 0.04$, $p = .518$, $\eta^2_g = .002$, was significant.

[2] Following each true/false judgment made during the test phase, we also included an exploratory task in which participants judged the temporal placement of the tested event. A solid white line 20 pixels in height was used for this purpose, appearing below the true/false statement after each true/false judgment. This line represented the timeline of the tested video and disappeared once the participant used the computer mouse to indicate where along the line the event occurred. The purpose of this task was to determine whether intentional forgetting affected the temporal specificity of the to-be-forgotten events. Each judgment was converted into a proportion representing the amount of the video preceding that point. The absolute difference between this value and the true position of the event within the video was calculated for true statements and analyzed as a function of instruction (R, F) and specificity (specific, general). Nothing was significant (all $p > .476$ and all $\eta^2_g = .001$). For F segments the mean deviation was 0.15 ($SE = 0.01$) for general statements and 0.15 ($SE = 0.01$) for specific statements; for R segments the mean deviation was 0.15 ($SE = 0.01$) for general statements and 0.15 ($SE = 0.01$) for specific statements. It appears that the task was simply too difficult. The magnitude of the absolute difference scores converts to an approximate granularity of 42 ms. This suggests that participants had only an approximate notion of the segment in which any given event occurred.

Table 2
Mean probe reaction times (in milliseconds) and associated accuracies (%) as a function of instruction and SOA for Experiments 4 and 5.

SOA	Reaction time (ms)				Accuracy (%)			
	R		F		R		F	
	M	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Experiment 4								
1800 ms	577	54	667	57	93.75	2.26	95.83	3.25
2600 ms	553	58	646	64	98.96	1.04	96.88	1.72
3400 ms	578	54	590	54	98.96	1.04	97.92	1.44
Experiment 5								
800 ms	471	23	576	34	98.61	0.97	96.53	1.46
1200 ms	441	28	537	39	97.22	1.32	96.53	1.77
2000 ms	440	23	489	39	99.31	0.69	98.61	0.97

16.2. Probe detection RTs

A probe-present detection trial was considered correct if the participant executed a response within 100 ms and 2000 ms of probe onset; as depicted in Table 2, correct detection responses did not differ as a function of instruction, $F(1, 23) = 0.05$, $MSe = 94.92$, $p = .833$, $\eta^2_g < .001$, or SOA, $F(1, 23) = 1.94$, $MSe = 96.05$, $p = .155$, $\eta^2_g = .029$, and these factors did not interact, $F(1, 23) = 0.86$, $MSe = 65.48$, $p = .429$, $\eta^2_g = .009$. Mean probe RTs for correct trials were analyzed as a function of instruction (R, F) and SOA (1800 ms, 2600 ms, 4200 ms) using a repeated-measures ANOVA. This analysis revealed probe RTs to be significantly longer following F instructions ($M = 634$ ms, $SE = 56$ ms) compared to R instructions ($M = 567$ ms, $SE = 54$ ms), $F(1, 23) = 15.25$, $MSe = 10621.57$, $p < .001$, $\eta^2_g = .015$. The main effect of SOA was not significant, $F(1, 23) = 1.54$, $MSe = 12940.20$, $p = .226$, $\eta^2_g = .004$. There was, however, a marginal instruction × SOA interaction, $F(1, 23) = 3.10$, $MSe = 7232.33$, $p = .055$, $\eta^2_g = .004$. Planned contrasts evaluated the difference between the F and R trials for each SOA. A significant 91 ms R > F difference was observed for the 1800 ms condition, $t(23) = 3.25$, $p = .003$, as well as a significant 93 ms R > F difference for the 2600 ms condition, $t(23) = 2.81$, $p = .009$. However, the 17 ms R > F difference observed for the 4200 ms condition failed to reach significance, $t(23) = 1.15$, $p = .261$.

16.3. Probe false alarms

The percentage of false alarms was analyzed as a function of instruction (R, F) revealing a significant difference, $F(1, 23) = 7.67$, $MSe = 61.14$, $p = .011$, $\eta^2_g = .250$. Participants did not commit any false alarms (i.e., 0.00%) for F trials whereas for R trials on average 6.25% ($SE = 2.26\%$) of catch trials resulted in a false alarm. Although congruent with a similar finding by Fawcett and Taylor (2008), this particular analysis is difficult to interpret due to the extreme floor effect resulting in no variability for the F trials.

17. Discussion

Once again a significant directed forgetting effect was observed for specific but not general test statements as measured by A'. Further, a time-course was revealed for which the F > R probe RT difference was largest at 2600 ms and dissipated by 4200 ms following instruction onset. This finding contrasts with the time-course observed by Fawcett and Taylor (2008) in which the F > R probe RT difference was most robust at 1800 ms and dissipated by 2600 ms following instruction onset. While it could be that the time-course differed simply because we employed a novel range of SOAs (see Cheal & Chastain, 2002), it is also possible that the more complex nature of our stimuli (segments of continuous visual events) makes them more difficult to intentionally forget than the words used by Fawcett and Taylor (2008). As a result, participants could spend longer implementing the memory instructions, resulting in the observed time-course.

18. Experiment 5

Experiment 5 replicated the methods of Experiments 2 and 4 with the exception that specificity was manipulated between- as opposed to within-subjects. In Experiments 2 and 4, there was necessary overlap between the information tested by some of the specific statements (e.g., *the woman added 3 cups of flour*) and some of the general test statements (e.g., *the woman added flour*). In cases where there was conceptual overlap between the general and specific statements, it seems possible that exposure to one type of statement might influence subsequent performance on the other; it seems particularly likely that exposure to the specific test statement might provide sufficient information to influence the response made to subsequent presentation of the conceptually overlapping general test statement. Demonstrating a comparable pattern of results, with a directed forgetting effect for specific but not for general statements in a between-subjects manipulation would eliminate this concern. Experiment 5 also further explored the time-course revealed in Experiments 2 and 4 by presenting probes at 800 ms, 1200 ms and 2000 ms instruction-probe SOAs.

19. Method

19.1. Participants

Thirty-six undergraduate students (31 female) enrolled at Dalhousie University participated in this experiment for course credit. Most participants were right-handed (32 right). Half of these participants were randomly assigned to receive the specific test statements and the remaining half of these participants were randomly assigned to receive the general test statements.

19.2. Stimuli and apparatus

The stimuli and apparatus were identical to Experiment 2.

19.3. Procedure

The procedure was identical to Experiment 2 with the following exceptions. First, during the study phase, visual probes were presented 800 ms, 1200 ms or 2000 ms following memory instruction onset. These particular SOAs were selected to more fully explore the time-course in close temporal proximity to the memory instruction. The timing of the study phase trials otherwise remained unchanged. Second, during the test phase either the specific *or* the general test statements were presented and this variable was manipulated between-subjects.

20. Results

20.1. Signal detection analysis

The mean hits and false alarms are provided in Table 1. Non-parametric measures of sensitivity (A') and response bias (B''_D) were calculated and analyzed as a function of instruction (R, F) and relative specificity (specific, general) using separate two-way mixed ANOVAs (see Fig. 2). For the A' analysis, there was a significant effect of instruction with greater sensitivity to statements about R segments than F segments, $F(1, 34) = 4.63$, $MSe = 0.002$, $p = .039$, $\eta_g^2 = .028$. The main effect of specificity was also significant with greater sensitivity to general statements than specific statements, $F(1, 34) = 31.18$, $MSe = 0.008$, $p < .001$, $\eta_g^2 = .419$. Importantly, these effects were qualified by a significant instruction × specificity interaction, $F(1, 34) = 5.27$, $MSe = 0.002$, $p = .028$, $\eta_g^2 = .032$. Planned contrasts revealed a significant 0.04 directed forgetting effect in A' for specific test statements, $t(17) = 2.66$, $p = .017$. There was no difference in A' measured on R and F trials for general test statements, $t(17) = 0.13$, $p = .897$.

The B''_D analysis revealed only a significant main effect of relative specificity, $F(1, 34) = 6.94$, $MSe = 0.11$, $p = .013$, $\eta_g^2 = .116$, with participants employing a more liberal response bias for general statements than specific statements; neither the main effect of instruction, $F(1, 34) = 1.58$, $MSe = 0.06$, $p = .217$, $\eta_g^2 = .016$, nor the instruction × specificity interaction, $F(1, 34) = 1.02$, $MSe = 0.006$, $p = .319$, $\eta_g^2 = .011$, was significant.

20.2. Probe detection RTs

The specific and general groups were collapsed for the purpose of this analysis (as well as the following analyses) because the study phase was identical for these groups. A probe-present detection trial was considered correct if the participant executed a response within 100 ms and 2000 ms of probe onset; as depicted in Table 2, correct detection responses did not differ as a function of instruction, $F(1, 35) = 1.70$, $MSe = 42.58$, $p = .201$, $\eta_g^2 = .006$, or SOA, $F(2, 70) = 1.29$, $MSe = 63.16$, $p = .284$, $\eta_g^2 = .014$, and these factors did not interact, $F(2, 70) = 2.17$, $MSe = 53.24$, $p = .805$, $\eta_g^2 = .002$. Mean probe RTs for correct trials were analyzed as a function of instruction (R, F) and SOA (800 ms, 1200 ms, 2000 ms) using a repeated-measures ANOVA. As in the preceding experiments, probe RTs were significantly longer following F instructions ($M = 534$ ms, $SE = 35$ ms) compared to R instructions ($M = 451$ ms, $SE = 22$ ms), $F(1,35) = 18.60$, $MSe = 20059.00$, $p < .001$, $\eta_g^2 = .047$. The main effect of SOA was not significant, $F(2, 70) = 6.76$, $MSe = 9390.36$, $p = .002$, $\eta_g^2 = .016$, nor was the instruction × SOA interaction, $F(2, 70) = 2.22$, $MSe = 7304.73$, $p = .116$, $\eta_g^2 = .004$. Despite the non-significant interaction, planned contrasts evaluated the difference between F and R trials for each SOA. A significant 105 ms difference was observed for the 800 ms condition, $t(35) = 4.21$, $p < .001$; a significant 96 ms difference was observed for the 1200 ms condition, $t(35) = 3.76$, $p < .001$; and, a marginally significant 49 ms difference was observed for the 2000 ms condition, $t(35) = 1.91$, $p = .065$.

20.3. Probe false alarms

The percentage of false alarms was analyzed as a function of instruction (R, F) using a repeated-measures ANOVA. Although fewer false alarms were committed following F instructions ($M = 4.17\%$, $SE = 1.86\%$) compared to following R instructions ($M = 8.33\%$, $SE = 2.23\%$), this difference was not significant, $F(1, 35) = 2.06$, $MSe = 151.7857$, $p = .160$, $\eta_g^2 = .056$.

21. Discussion

Experiment 5 revealed a directed forgetting effect for specific but not general test statements, even though specificity was manipulated between subjects. This precludes the possibility that exposure to specific test statements unduly influenced performance for the general test statements (or vice versa) in our preceding experiments. Experiment 5 also explored much earlier instruction-probe SOAs than previous experiments using simple detection. The time-course for the probe RTs in the current experiment was generally consistent with the time-course observed for the preceding experiments.

22. General discussion

Five experiments investigated the intentional forgetting of visual vignettes in a paradigm that presented an R or F instruction following discrete video segments that otherwise comprised a continuous event sequence. Participants demonstrated superior memory performance for R segments compared to F segments when tested using relatively specific questions or true/false statements but not when tested using relatively general true/false statements. In other words, there was a directed forgetting effect for specific information that occurred despite the fact that participants had no basis for predicting a priori whether a given video segment would be followed by an R or F instruction; pre-instruction encoding of R and F segments was thus equated, leaving only post-instruction processes to account for the directed forgetting effect observed for the specific details depicted in those segments.

The notion that intentional forgetting may affect memory in a graded fashion – with specific details being more affected by memory instruction than general details – speaks to the representation of the to-be-forgotten information in memory. If intentional forgetting was viewed as an all-or-nothing process then the unwanted F-instructed information should either be accurately reported on a subsequent memory test despite the intention to forget (unintentional remembering) or it should not be reported according to the formulated intention (intentional forgetting). Instead, the current results suggest that although the details of a to-be-forgotten memory are lost, a more general representation of the memory persists. This conclusion is supported by evidence that the to-be-forgotten information also persists within implicit memory (e.g., McKinney & Woodward, 2004), may influence subsequent social judgments (Bodenhausen, Macrae, & Milne, 1998; Golding & Hauselt, 1994) or jury decisions (for a review, see Kassin & Studebaker, 1998 or Steblay, Hosch, Culhane, & McWethy, 2006) and may under certain circumstances drive false recall or recognition of information with a similar gist (Kimball & Bjork, 2002).

Past studies using structured study materials such as semantically related word lists (e.g., Golding, Long, & Macleod, 1994) or textual materials (e.g., Johnson, 1998; Geiselman, 1974, 1977; see also, Delaney, Nghiem, & Waldum, 2009) have revealed that connections between the R and F information mitigate directed forgetting. It may be the case that the efficacy of an instruction to intentionally forget depends, at least in part, on whether the to-be-forgotten information is deemed both nominally and functionally irrelevant to the ongoing task. That is, even if the information is nominally tagged as irrelevant by virtue of receiving an F instruction, it may be functionally relevant for building an understanding of (and possibly for more easily retaining) the to-be-remembered information. This argument harkens back to Golding and Keenan (1985) who observed that participants retained spatial directions that had been labeled as erroneous to ensure that this otherwise irrelevant information did not lead them to make an incorrect navigational decision when enacting those directions. Of course, this view is post-hoc and is more relevant to how participants process the memory instruction as opposed to the cognitive or neural processes involved.

Of course, there are many possible interpretations regarding why we observe directed forgetting for specific but not general test statements. Perhaps re-exposure to the test statements relevant to the F segments resulted in a release of inhibition masking directed forgetting for general test statements and reducing the magnitude of directed forgetting for specific test statements (e.g., Bjork & Bjork, 1996). Unfortunately, this possibility is not easily addressed in the context of the current analyses. One difficulty with such an account would be the fact that directed forgetting is observed even for specific test statements: There is no a priori reason to suspect that a release of inhibition would preferentially benefit general test statements unless one also assumed that the magnitude of directed forgetting was smaller for general test statements. We are not ruling out this possibility and it is an avenue worth pursuing in future research. Further, our stimuli were not perfectly suited for our purposes (we used pre-existing, publically available videos) and differed from video to video and even from segment to segment in terms of their relative density of information. As a result, it could be that directed forgetting for general test statements has been "washed out" by videos or segments containing relatively few events or even events that spanned multiple segments. While inspection of the data suggests a similar pattern across each video, we cannot rule out this possibility. Further studies are required using custom-made stimuli with greater control over the nature and amount of testable information within each segment. Nonetheless, the current findings represent an important step in this direction and will hopefully stimulate further investigation of both intentional forgetting within event-memory and differences in how these effects apply to specific and general information.[3]

Experiments 1, 2, 4 and 5 also included a visual probe following the R or F memory instructions to gauge the cognitive load required to instantiate each instruction. Whereas proponents of a purely selective rehearsal based interpretation of the current findings might have predicted slower detection responses following R than F instructions, a growing body of literature within the item-method (e.g., Fawcett & Taylor, 2008; Hansen, 2011) has emerged demonstrating the opposite pattern: Participants are reliably slower to respond following F than R instructions within approximately 2600 ms of instruction onset. In the current paradigm, Experiments 1 and 2 measured detection RTs at 1800 ms following the onset of the memory instruction, whereas Experiments 4 and 5 manipulated the instruction-probe SOA to include 800 ms, 1200 ms, 2000 ms, 2600 ms and 4200 ms intervals as well. The time-course revealed by the combination of these experiments is plotted in Fig. 3 along with a regression line calculated by fitting a linear-mixed effects model to the mean F–R RT difference using SOA as a fixed-effect and participant and experiment as random effects (see Bates, 2007). The F > R pattern of RTs was evident as early as 800 ms although it diminished in magnitude as the trial progressed. This finding is similar to Fawcett and Taylor (2008) who observed that the F > R RT difference dissipated within 2600 ms of memory instruction onset. The relative protraction of the time-course beyond an SOA of 2600 ms in Fig. 3 may be attributable to differences in the effort required to intentionally forget isolated words or pictures (as in Fawcett & Taylor, 2008) as compared to the more complex event sequences used in the current study (see also Fawcett et al., in press). Nowicka, Marchewka, Jednoróg, Tacikowski, and Brechmann (2011) recently observed greater and more widespread neural activation associated with successfully forgetting emotionally salient rather than neutral pictures. They argued that forgetting emotional memories requires greater effort. To the degree that the neural activation measured by Nowicka et al. (2011) and the probe RTs measured in the current study index a common pool of cognitive processes that are marshaled by an intention to forget, their argument could be extended to converge upon the hypothesis that the effort required to intentionally forget a memory is proportional to the salience or contextual "richness" of the to-be-forgotten trace.

The current study fits into a much larger literature providing evidence of an effortful mechanism associated with intentional forgetting. Fawcett and Taylor (2012) found that in addition to slowing discrimination responses, instantiating an item-method F instruction also interfered with incidental memory formation. This is in addition to the work conducted by Taylor (2005) (see also, Fawcett & Taylor, 2012; Taylor & Fawcett, 2011) who demonstrated a larger inhibition of return effect following F instructions rather than R instructions. Taylor and Fawcett (2011) speculated that enacting an F instruction might lead the participant to respond more cautiously towards subsequent information arising from the source of the to-be-forgotten

[3] We thank two anonymous reviewers for raising the issues discussed in this paragraph.

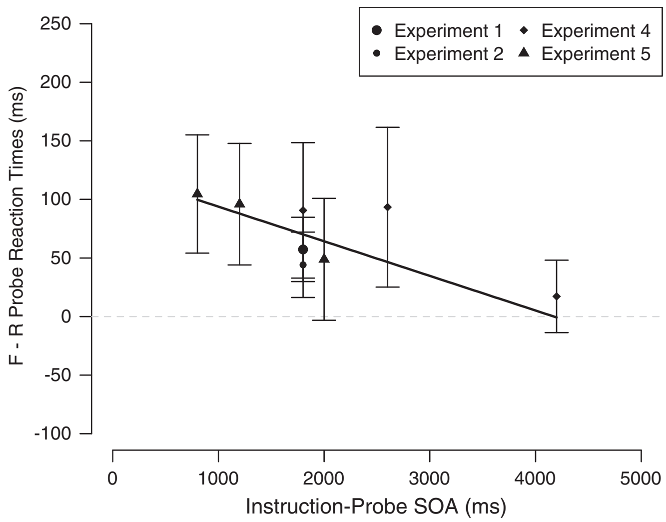

Fig. 3. Time-course of the F > R probe reaction time difference for Experiments 1, 2, 4 and 5; error bars represent the 95% confidence interval of the mean difference and the regression line was calculated using a linear-mixed effects regression model with SOA as a fixed effect and participant and experiment as random effects (Bates, 2007).

information. Indeed, Fawcett and Taylor (2010) observed that participants were better at countermanding an unwanted motor response in a stop-signal task when the stop-signal was preceded by an F instruction rather than an R instruction: Instantiating the F instruction slowed subsequent responding, which in turn increased the likelihood that the participant would be able to use the stop signal to successfully cancel the prepotent motor response. Together with recent neuroimaging studies (e.g., Nowicka et al., 2011; Paz-Caballero, Menor, & Jimenez, 2004; Wylie et al., 2008) these findings implicate an active mechanism that discourages the continued processing of unwanted information.

One possible mechanism proposed to account for the F > R effect using the item-method paradigm has been the retrieval and cumulative rehearsal of the preceding to-be-remembered materials following each F instruction. Fawcett and Taylor (2008) (see also, Fawcett & Taylor, 2010, 2012) discussed and rejected this notion in the context of their original item-method probe studies. Indeed, when participants are asked to describe the strategies they employed during the study phase of a directed forgetting task they will often include cumulative rehearsal of the preceding R items among other strategies. According to this view, the pattern of slower probe RTs following F compared to R instructions is attributable to the effortful search required to retrieve the prior R items. While Fawcett and Taylor (2008) (see also, Fawcett & Taylor, 2010, 2012) did not dispute that participants may indeed engage in cumulative rehearsal across trials, they discussed several reasons why this strategy was not primarily responsible for their F > R RT probe findings. Nevertheless, one might question whether cumulative rehearsal could account for the current F > R RT difference for responding to the probe.

One way to address the concern of cumulative rehearsal would be to demonstrate the presence of an F > R probe RT difference under circumstances in which cumulative rehearsal was unlikely to occur. We reasoned that cumulative rehearsal was likely to occur within but not between videos. According to this reasoning, the initial segment of each video offered a scenario in which participants were unlikely to retrieve or rehearse preceding R segments — allowing measurement of the F > R probe RT difference without contamination from retrieval processes. We collapsed the probe RTs from the initial segment of each video for Experiments 1, 2, 4 and 5 and analyzed them as a function of instruction (R, F) using a linear-mixed effects model with participant and video as random effects (Bates, 2007). We had initially included presentation order (1, 2, 3, 4) in our model but this factor was removed because neither it, $ß = 10.90$, $t = 0.836$, $SE = 13.03$, nor the instruction × order interaction, $ß = -6.68$, $t = -0.32$, $SE = 20.80$, was significant. There was, however, a significant 70 ms F > R RT difference, $ß = -69.63$, $t = 2.95$, $SE = 23.60$, which was of consistent magnitude across presentation order. This finding provides compelling evidence that the active mechanism(s) indexed following each F instruction occur even when cumulative rehearsal is improbable or even impossible.[4]

Although it remains to be seen whether the mechanisms identified in the current event-method paradigm are indeed comparable to those previously identified in an item-method task, the current experiments provide evidence that in the context of the delayed presentation of each memory instruction this may be the case. Fawcett et al. (in press) argued that when the memory instruction was presented concurrent with the studied materials that instead of controlling the *contents* of working memory resources, participants must instead control *access* to working memory resources. It is probable that this is just as true when dealing with pictures and words as when dealing with events. The consequence of intentional forgetting appears to be a relatively impoverished memory trace for the to-be-forgotten materials in relation to the to-be-remembered materials. This robust directed forgetting effect for specific but not for general details is true whether the R and F instructions are presented concurrent with or following the studied materials and is attributable to the costs rather than the benefits of intentional forgetting.

Author Note

We would like to thank Thanh Luu for his help in collecting data, Jonathan Crowell for his help in coding the data, Siobhan Hoscheidt for her assistance with the ethical review process, Dr. John Christie for his statistical advice and Dr. Ray Klein and Dr. Vin LoLordo for their feedback and guidance. JMF would further like to thank the University of Arizona for hosting him throughout this project where he was funded by an NSERC Canadian Graduate Scholarship, Killam Predoctoral Scholarship and Michael Smith Foreign Study Supplement. This research was also supported by a funding from NSERC to TLT.

References

Basden, B. H., & Basden, D. R. (1996). Directed forgetting: Further comparisons of the item and list methods. *Memory*, *4*(6), 633–653.
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., & Gargano, G. J. (1993). Directed forgetting in implicit and explicit memory tests: A comparison of methods. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 603–616.
Bates, D. (2007). *lme4: Linear mixed-effects models using S4 classes.* R package version 0.99875-9.
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.

[4] We recognize that the ideal analysis would be to demonstrate the F > R RT difference for only the first trial as opposed to the first segment of each video. Such an analysis relies upon a very small sample of data – a single RT from each subject – resulting in especially low statistical power, especially for a between-subjects analysis. There are two reasons that we feel our first segment analysis is still a suitable response to the notion of cumulative rehearsal. First, as discussed in-text we initially included presentation order as a factor and found that neither its main effect nor the instruction × presentation order interaction reached significance. That is to say that the F > R RT difference was of similar magnitude for the first segment of the first video as it is for the first segment of the second, third and fourth videos. Second, an analysis which pools data from several experiments (including those in the current study) reveals a significant R > F RT difference for the initial trial (Fawcett, Taylor & Nadel, 2011). Combined with the breadth of evidence concerning this active mechanism (e.g., Fawcett & Taylor, 2008; Taylor & Fawcett, 2011) the cumulative rehearsal of preceding R information on F trials cannot provide a complete or compelling explanation of the F > R RT difference.

Bjork, E. L., & Bjork, R. A. (1996). Continuing influences of to-be-forgotten information. *Consciousness and Cognition*, *5*, 176–196.

Bjork, R. A., & Woodward, A. E. (1973). Directed forgetting of individual words in free recall. *Journal of Experimental Psychology*, *99*, 22–27.

Bodenhausen, G. V., Macrae, C. N., & Milne, A. B. (1998). Disregarding social stereotypes: Implications for memory, judgment and behavior. In J. M. Golding, & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 349–369). Mahwah, NJ: Erlbaum.

Cheal, M., & Chastain, G. (2002). Timing of facilitatory and inhibitory effects of visual attention. *Visual Cognition*, *9*(8), 969–1002.

Delaney, P. F., Nghiem, K. N., & Waldum, E. R. (2009). The selective directed forgetting effect: Can people forget only part of a text? *Quarterly Journal of Experimental Psychology*, *62*(8), 1542–1550.

Donaldson, W. (1992). Measuring recognition memory. *Journal of Experimental Psychology. General*, *121*(3), 275–277.

Fawcett, J. M., Lawrence, M. A., & Taylor, T. L. (2011). *Item-method directed forgetting is effortful and impoverishes memory for abstract images.* Poster presented at the annual meeting of the Society for Object Perception, Attention and Memory, Seattle, Washington.

Fawcett, J. M., & Taylor, T. L. (2008). Forgetting is effortful: Evidence from reaction time probes in an item-method directed forgetting task. *Memory and Cognition*, *36*(6), 1168–1181.

Fawcett, J. M., & Taylor, T. L. (2010). Directed forgetting shares mechanisms with attentional withdrawal but not with stop-signal inhibition. *Memory and Cognition*, *38*(6), 797–808.

Fawcett, J. M., & Taylor, T. L. (2012). The control of working memory resources in intentional forgetting: Evidence from incidental probe word recognition. *Acta Psychologica*, *139*, 84–90.

Fawcett, J. M., Taylor, T. L., & Nadel, L. (2011). *Forgetting is more effortful than remembering: A first-trial analysis of post-instruction probes in an item-method task.* Unpublished manuscript.

Fawcett, J. M., Taylor, T. L., & Nadel, L. (2013). Intentional forgetting diminishes memory for continuous events. *Memory*, http://dx.doi.org/10.1080/09658211.2012.748078 (in press).

Geiselman, R. E. (1974). Positive forgetting of sentence material. *Memory and Cognition*, *2*, 677–682.

Geiselman, R. E. (1977). Memory for prose as a function of learning strategy and inspection time. *Journal of Educational Psychology*, *69*, 547–555.

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., & Hauselt, J. (1994). When instructions to forget become instructions to remember. *Personality and Social Psychology Bulletin*, *20*(2), 178–183.

Golding, J. M., & Keenan, J. M. (1985). Directed forgetting and memory for directions to a destination. *The American Journal of Psychology*, *98*(4), 579–590.

Golding, J. M., Long, D. L., & Macleod, C. M. (1994). You can't always forget what you want: Directed forgetting of related words. *Journal of Memory and Language*, *33*, 493–510.

Gottlob, L. R., Golding, J. M., & Hauselt, W. J. (2006). Directed forgetting of a single item. *The Journal of General Psychology*, *133*(1), 67–80.

Hansen, W. (2011). The reality of directed forgetting in the item-method paradigm: Suppression, not selective search or decay. Doctoral Dissertation. Retrieved on March 10 2012, from. http://hdl.handle.net/2286/az5d5178n18

Hourihan, K. L., Ozubko, J. D., & MacLeod, C. M. (2009). Directed forgetting of visual symbols: Evidence for nonverbal selective rehearsal. *Memory and Cognition*, *37*(8), 1059–1068.

Johnson, H. M. (1998). Disregarding information in text. In J. M. Golding, & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 219–239). Mahwah, NJ: Erlbaum.

Joslyn, S. L., & Oakes, M. a (2005). Directed forgetting of autobiographical events. *Memory and Cognition*, *33*(4), 577–587.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

Kassin, S. M., & Studebaker, C. A. (1998). Instructions to disregard the jury: Curative and paradoxical effects. In J. M. Golding, & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 413–435). Mahwah, NJ: Erlbaum.

Kimball, D. R., & Bjork, R. A. (2002). The influence of intentional and unintentional forgetting on false memories. *Journal of Experimental Psychology. General*, *131*, 116–130.

MacLeod, C. M. (1989). Directed forgetting affects both direct and indirect tests of memory. *Journal of Experimental Psychology: Learning, Memory, and 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.

McKinney, L. C., & Woodward, A. E. (2004). Remembering what one intended to remembering forgetting forget: The lack of directed effects in implicit memory. *The American Journal of Psychology*, *117*(2), 169–190.

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 and Emotion*, *18*, 479–493.

Nowicka, A., Marchewka, A., Jednoróg, K., Tacikowski, P., & Brechmann, A. (2011). Forgetting of emotional information is hard: An fMRI study of directed forgetting. *Cerebral Cortex*, *21*(3), 539–549.

Pastotter, B., Kliegl, O., & Bauml, K. -H. (2012). List-method directed forgetting: The forget cue improves both encoding and retrieval of postcue information. *Memory and Cognition*, *40*, 861–873.

Paz-Caballero, M. D., Menor, J., & Jimenez, J. M. (2004). Predictive validity of event-related potentials (ERPs) in relation to the directed forgetting effects. *Clinical Neurophysiology*, *115*, 369–377.

Quinlan, C. K., Taylor, T. L., & Fawcett, J. M. (2010). Directed forgetting: Comparing pictures and words. *Canadian Journal of Experimental Psychology*, *64*(1), 41–46.

Racsmány, M., & Conway, M. A. (2006). Episodic inhibition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(1), 44–57, http://dx.doi.org/10.1037/0278-7393.32.1.44.

Sahakyan, L., & Foster, N. L. (2009). Intentional forgetting of actions: Comparison of list-method and item-method directed forgetting. *Journal of Memory and Language*, *61*(1), 134–152.

Sahakyan, L., & Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*(6), 1064–1072.

Steblay, N., Hosch, H. M., Culhane, S. E., & McWethy, A. (2006). The impact on juror verdicts of judicial instruction to disregard inadmissible evidence: A meta-analysis. *Law and Human Behavior*, *30*(4), 469–492.

Taylor, T. (2005). Inhibition of return following instructions to remember and forget. *The Quarterly Journal of Experimental Psychology: Section A*, *58*(4), 613–629.

Taylor, T. L., & Fawcett, J. M. (2011). Larger IOR effects following forget than following remember instructions depend on exogenous attentional withdrawal and target localization. *Attention, Perception, & Psychophysics*, *73*(6), 1790–1814.

Wylie, G. R., Foxe, J. J., & Taylor, T. L. (2008). Forgetting as an active process: An fMRI investigation of item-method-directed forgetting. *Cerebral Cortex*, *18*(3), 670–682.

Zacks, J. M., Tversky, B., & Iyer, G. (2001). Perceiving, remembering, and communicating structure in events. *Journal of Experimental Psychology. General*, *130*(1), 29–58.