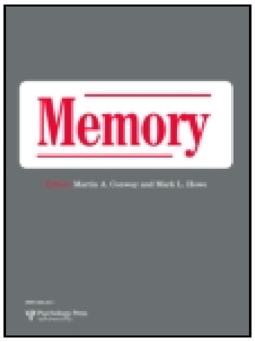

Memory

ISSN: 0965-8211 (Print) 1464-0686 (Online) Journal homepage: https://www.tandfonline.com/loi/pmem20

Intentional forgetting diminishes memory for continuous events

Jonathan M. Fawcett , Tracy L. Taylor & Lynn Nadel

Memory, 2013
Vol. 21, No. 6, 675–694, http://dx.doi.org/10.1080/09658211.2012.748078

Routledge
Taylor & Francis Group

Intentional forgetting diminishes memory for continuous events

Jonathan M. Fawcett[1], Tracy L. Taylor[1], and Lynn Nadel[2]

[1]Department of Psychology, Dalhousie University, Halifax, Nova Scotia, Canada
[2]Department of Psychology, University of Arizona, Tucson, AZ, USA

In a novel *event method* directed forgetting task, instructions to Remember (R) or Forget (F) were integrated throughout the presentation of four videos depicting common events (e.g., baking cookies). Participants responded more accurately to cued recall questions (E1) and true/false statements (E2-4) regarding R segments than F segments. This was true even when forced to attend to F segments by virtue of having to perform concurrent discrimination (E2) or conceptual segmentation (E3) tasks. The final experiment (E5) demonstrated a larger R > F difference for specific true/false statements (*the woman added three cups of flour*) than for general true/false statements (*the woman added flour*) suggesting that participants likely encoded and retained at least a general representation of the events they had intended to forget, even though this representation was not as specific as the representation of events they had intended to remember.

Keywords: Intentional forgetting; Memory; Cognition; Events.

Memory research tends to focus on those factors that contribute to the intentional *remembering* of information, often neglecting the factors that contribute to the intentional *forgetting* of information. This fixation on the processes associated with remembering as opposed to forgetting is unfortunate, because forgetting is essential to the efficiency of remembering. Forgetting encourages the dissolution of information that is no longer relevant and which could otherwise interfere with information that remains relevant (e.g., Bjork, 1972). The importance of forgetting to efficient mnemonic function was recognised as early as Cicero (1840/2009) and again by William James (1890/1950) but only in recent decades has it received experimental attention commensurate with its role in cognition.

Since the seminal work of Muther (1965) and later Bjork, LaBerge, and Legrand (1968), intentional forgetting has been increasingly studied in the laboratory using a paradigm known as directed forgetting (for a review, see MacLeod, 1998). There have been many variants of the directed forgetting paradigm since its inception (for a review see Basden & Basden, 1998; Bjork, 1972), most of which have been categorised as belonging to either the item method or the list method. In an item method task study items are presented one at a time, each accompanied or followed by an instruction to remember (R) or

Address correspondence to: Jonathan M. Fawcett, Dalhousie University, Department of Psychology, Halifax, NS, CANADA, B3H 4J1. E-mail: jmfawcet@dal.ca

We would like to thank Thanh Luu for his help collecting data and Siobhan Hoscheidt for her assistance with the ethical review process. We would also like to thank PsycheTruth for permission to include screen shots of their video ‘‘How To Clean A Fish Tank or Aquarium, Remove Algae, Pet Care’’ (referred to in-text as ‘‘Cleaning a Fish Tank’’). JMF would further like to thank Dr Lynn Nadel and the University of Arizona for hosting him throughout this project during which he was funded by an NSERC Canadian Graduate Scholarship with a Michael Smith Foreign Study Supplement and then by a Killam Predoctoral Scholarship. This research was supported by funding from NSERC to TLT.

forget (F). In a list method task a single R or F instruction is presented following a discrete list of items, after which participants are asked to remember a second list. In both paradigms participants are subsequently tested for all study items, regardless of the R or F memory instruction. One common finding, referred to as a directed forgetting effect, is better memory performance for R items than for F items.

The nomenclature used to describe the experimental techniques with which directed forgetting is frequently studied has become synonymous with their respective methodological features. That is to say that list method tasks involve the intentional forgetting of lists whereas item method tasks involve the intentional forgetting of individual items. The relation between these procedures and the ability to intentionally forget more typical experiences is obscured by the fact that most episodes are not easily rendered as either lists or individual items. More frequently our memories involve sequences of experienced or performed events or actions—a fact evident in the examples commonly used to describe intentional forgetting (e.g., Bjork, 1972). Even so, studies using both the list and item method paradigms have often used pictures (e.g., Quinlan, Taylor, & Fawcett, 2010) or words (e.g., Tekcan & Aktürk, 2001) as the to-be-remembered or to-be-forgotten stimuli. Few studies have addressed the intentional forgetting of continuous episodes. Those that do often identify with the list method framework even though it is not clear that the same mechanisms used to forget lists of discrete items are necessarily applied to sequences of experienced events.

Earles and Kersten (2002) presented a series of verb–noun pairs describing an action (e.g., *break toothpick*) that was either performed or simply studied in a between-participants design. Following each verb–noun pair participants were asked to either remember or forget the action with the hypothesis that performed actions would be more difficult to intentionally forget than read actions. Even though older adults did not show the predicted pattern, younger participants exhibited a smaller directed forgetting effect for actions they had performed relative to those they had only studied. More recently, Sahakyan and Foster (2009) replicated this finding using both the item method (Experiment 4) and the list method (Experiments 1–3). They observed that a list method instruction to forget *studied* verb–noun pairs resulted in worse memory for List 1 F items and better memory for List 2 R items compared to a remember-all baseline group; however, a list method instruction to forget *performed* verb–noun pairs resulted only in worse memory for List 1 F items (not better memory for List 2 R items). Taken together, the studies by Earles and Kersten (2002) and Sahakyan and Foster (2009) provide critical evidence that directed forgetting may be applied to simple actions performed in isolation—however, the relevance of their findings to complex or observed actions requires elaboration (see also Burwitz, 1974).

Towards this goal, Joslyn and Oakes (2005) had participants record personally experienced events throughout a 2-week period with an emphasis on events atypical to their routine. Presented as a real-world analogue of the list method paradigm, half of these participants were later instructed that they would not be tested for (and therefore could forget) events occurring in the first week of the experiment (as recorded in their diary). The remaining participants were instructed they were required to remember these events in addition to any events occurring during the second week. A directed forgetting effect was observed in the recall of a two-word descriptive title (e.g., *Shopping Trip*) provided for each event as part of the diary exercise: Participants who received the F instruction remembered fewer Week 1 events than participants who did not receive this instruction.

While innovative, Joslyn and Oakes' (2005) methodology suffers from a lack of control inherent in any manipulation occurring outside the laboratory. Therefore Barnier et al., (2007) investigated the ability to intentionally forget experienced events—this time in a more controlled environment. Participants were presented with two lists each containing several cue words intended to elicit an autobiographical memory (e.g., *university*); once retrieved, the memory and the cue word were recorded. Once again branded as a list method task, half of the participants were instructed to forget the memories they had described in response to the first list of cue words prior to receiving the second list. Participants who received the F instruction later recalled fewer List 1 memories than those who did not receive this instruction.

The experiments described above represent an important step towards the application of intentional forgetting to the control of event memory—however, they have made this step within the confines of the item method and list method paradigms: In all cases participants were instructed to remember or forget a list of discrete events or actions—as provided in the form of verb–noun pairs (Earles & Kersten, 2002; Sahakyan & Foster, 2009) or as recorded in a diary (Barnier et al., 2007; Josyln & Oakes, 2005). As a result the subsequent tests focused on memory for the event more generally (e.g., *Did I break a match?*) as opposed to the details of that event (e.g., *What colour was the match that I broke?*). Moreover, when a diary or journal method is used, participants are required to provide the initial content that is used later in the test of memory (e.g., Barnier et al., 2007; Joslyn & Oakes, 2005): Sampling bias would seem to favour the selection of easily remembered central or salient details (as opposed to peripheral details) for initial report and subsequent inclusion in the memory test (see Joslyn & Oakes, 2005).

One method of circumventing these issues and advancing our understanding of the manner in which events may be intentionally forgotten is to investigate the ability to intentionally forget continuous events presented in the laboratory under controlled conditions. It is perhaps surprising that no published studies have yet utilised the potential for videos to accomplish these goals. Importantly, videos need not be separated into lists of discrete events followed individually (item method) or as a group (list method) by an instruction to remember or forget—instructions may instead be incorporated into the videos themselves such that segments of the otherwise continuous video events are associated with the R or F instructions.

THE CURRENT STUDY

We developed a novel *event method* directed forgetting paradigm that embeds memory instructions into continuous video sequences. Each video depicted a common event such as baking cookies and lasted for 4 minutes and 40 seconds. Because it was our desire to study how participants could selectively forget segments of *continuous* visual events we employed a *concurrent* memory instruction (e.g., Basden & Basden, 1996; Brown, 1954; Paller, 1990; Paller, Bozic, Ranganath, Grabowecky & Yamada, 1999; see also Muther, 1965) as opposed to a *delayed* memory instruction (for a review, see MacLeod 1998). Whereas a concurrent memory instruction is somewhat unusual in the directed forgetting literature because it risks having participants orient away from the input during presentation of the to-be-forgotten information, we considered this less of a risk in the current study. This is because our stimuli consisted of a continuous video event during which the memory instruction changed randomly between R and f instructions, such that participants could not predict when the next change would occur. This is different from a task where a word is presented individually accompanied by an instruction to remember or forget. In such a paradigm the participant could elect to orient away from the to-be-forgotten word and orient back when the next word was due. But when the memory instruction changes at unpredictable intervals during the continuous play of a single event sequence, orienting away would be counterproductive. Such a strategy would not only risk missing the next transition from an F to an R instruction, it would also undermine comprehension of the vignette and, in so doing, likely increase the difficulty of committing the R segments to memory. The benefit of presenting a concurrent memory instruction during the presentation of our video sequences is that it allowed us to present these sequences uninterrupted. Thus the video event could be conceived of as a single continuous event rather than a discrete series of related segments. The question, of course, is whether participants have the cognitive flexibility to selectively encode only portions of this otherwise continuous event sequence.

Videos were superimposed upon a coloured rectangular viewing area that was larger than the video presentation port (creating the appearance of a coloured border surrounding the video) and that periodically changed between green and purple. Participants were instructed that whenever the border surrounding the video was green they were to remember everything that was shown (R segment) because they would be tested for that information later; whenever the border surrounding the video was purple participants were instead to forget everything that was shown (F segment). Following the presentation of the study videos in each experiment, participants were tested for the details of the event using cued recall questions (Experiment 1) or true/false statements (Experiments 2–5). Experiments 1 and 2 demonstrated that participants were more accurate when tested for R segments than F segments, consistent with a directed forgetting effect. Experiments 3 and 4 replicated Experiment 2 while requiring participants to engage in a secondary task intended to ensure that visual attention was focused on the video (Experiment 3) or that encouraged conceptual processing of the video (Experiment 4) at all times—even during the presentation of the F instruction. Experiment 5 demonstrated that the directed forgetting effect observed in Experiment 2 was smaller for general as opposed to specific test statements.

The use of true/false test statements in Experiments 2–5 was particularly noteworthy. Because the nature of our paradigm permitted the creation of false statements for each individual R or F segment, separate false alarm rates (i.e., incorrect affirmation of a false statement) could be calculated for R and F conditions. As a result we were able to calculate measures of *sensitivity* (A') and *response bias* (B''_D) not frequently available in a typical item method or list method task where R and F conditions more typically share a common false alarm rate (see Zacks, Radvansky, & Hasher, 1996, Note 2).

EXPERIMENT 1

In Experiment 1 four videos were presented depicting common events such as baking cookies or preparing for work. For each video participants were instructed to remember a random half of its segments and to forget the remainder. The R and F instructions were integrated within each video and denoted by a change in the colour of the surrounding border (e.g., *remember anything that occurs when the border is green*). Thus each video played continuously from start to finish, with the surrounding border changing colour randomly throughout the presentation. Following the presentation of all four videos participants were then presented with a series of cued recall questions testing their memory for *all* video segments regardless of the previously associated memory instruction. We predicted that participants would respond more accurately when tested for segments they had been instructed to remember relative to those they had been instructed to forget. This finding would support our contention that directed forgetting occurs for continuous events and permit closer examination of this effect in the following experiments.

Method

Participants

A total of 30 undergraduate students (20 female) enrolled at the University of Arizona participated in this experiment for course credit. The majority of participants were right-handed (26 right, 4 left); their ages ranged from 18 to 25 years, with a mean of 18.77 years.

Stimuli and apparatus

All experimental procedures were presented using custom software developed in the Python programming language (www.python.org) with the Pygame development library (www.pygame.org) loaded on a 17-inch MacBook Pro computer running Mac OS X 10.5. Stimuli were viewed from an approximate distance of 57 cm and 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/). Each video presented during the study phase was preceded by a title that also served as a retrieval cue during the subsequent test phase; the titles were presented in white against a black background using size 30 of the Gentium Basic Bold font.

Five videos were downloaded from the public domain video sharing website YouTube (www.youtube.com) to serve as stimuli in this experiment: The first video (Folding Laundry) was used for practice, and the remaining four videos (Cleaning a Fish Tank, Baking Assorted Cookies, Making Chocolate Pudding, and Getting Ready for Work) were used during the study phase. Videos were selected on the basis of two criteria: (a) Their content was easily understood in the absence of the associated audio track; and (b) they contained a linear progression of events resulting in a predetermined, self-evident goal which could be explicated in a short, descriptive title (e.g., Cleaning a Fish Tank). Once downloaded, these videos were converted to MPEG-1 format, resized to 600×600 pixels and edited until they contained 7000 frames (1750 frames for the practice video). Presented at an average rate of 25 frames per second, each video lasted 4 minutes and 40 seconds (1 minute and 10 seconds for the practice video). During the video conversion process the audio track was removed from each video.

The coloured border that acted as the R or F memory instruction throughout the practice and study phases subtended 35 pixels and surrounded each video: A green border denoted an R segment and a purple border denoted an F segment. The specific shades of green and purple are denoted by the RGB values of (0,100,0) and (128,0,128), respectively, and were selected on the basis that these colours are easily discriminated even in the presence of abnormal colour perception. The assignment of green and purple to remember and forget was constant as it was believed that green was more easily associated with the process of remembering (e.g., like a green traffic light) and we did not want to load participants' working memories with the colour-instruction mapping[1]. Each segment consisted of 875 frames (35 seconds), resulting in eight segments per video (two segments for the practice video). R segments and F segments were assigned randomly on a participant-by-participant basis with the caveat that each video always contained 4 R segments and 4 F segments (1 R segment and 1 F segment for the practice video).

Cued recall questions were designed to test specific details revealed only during a single segment of each video. For example, *how many sticks of butter did the woman add to the mixing bowl?* Two of the videos (Baking Assorted Cookies, Making Chocolate Pudding) were tested with 3 cued recall questions per segment and the remaining two videos (Cleaning a Fish Tank, Getting Ready for Work) were tested with two cued recall questions per segment for a total of 80 cued recall questions across all four videos. The number of test statements per segment reflected differences in the ''richness'' of the videos themselves with regard to the number and testability of events.

Procedure

Participants were told that they would view four videos each depicting an event such as folding laundry during which they would be instructed to remember only some of the information presented. Participants were instructed that whenever the border surrounding the video was green they were to remember everything that was shown because they would be tested for that information later; whenever the border surrounding the video was purple they were to forget everything that was shown. Participants were notified that the colour of this border would change after various intervals and that it was important that they continue attending to the computer screen to ensure that they did not miss one of these changes. As represented in the software code, the border colour changed at regular intervals of 35 seconds (see below); however, because the R and F segments were randomly interspersed throughout each video, the border sometimes changed from green to green (or purple to purple) such that, from the participants' perspective, the duration of the instructions seemed variable.

Practice phase. A practice video (Folding Laundry) was presented to familiarise participants with the task, during which the experimenter offered sample questions pertaining to the practice video so that the participant would understand the type of information they were expected to retain. The practice video lasted 70 seconds and was comprised of a single R segment and a single F segment lasting 35 seconds each. Once the practice video was over, participants were presented with a written version of the study phase instructions on the computer screen and told to press ENTER when they were ready to begin the experiment proper. The experimenter relocated to a different desk at the far end of the room behind the participant.

Study phase. Prior to each video a descriptive title (e.g., ''Cleaning a Fish Tank'') was presented in the centre of the screen until the participant pressed the ENTER key, at which point the video began. Videos were presented at an approximate rate of 25 frames per second (40 ms per frame) until all 7000 frames were exhausted. Videos were separated into eight segments each lasting 875 frames (35 seconds) during which the border surrounding the video was either green (for an R segment) or purple (for an F segment). Prior to beginning each video, half of these segments were designated as R segments and the remaining were designated as F segments. The assignment of R and F segments was randomised for each video on a participant-by-participant basis, as was the presentation order of the videos themselves. Once a given video ended, the title for the next video was presented (until the participant pressed the ENTER key). This process continued until the participant had watched all four videos, at which point the study phase ended and the written instructions for the test phase were displayed. Figure 1 provides a schematic representation of the study phase.

[1] This notion that the green colour was easily mapped to the R instruction and the purple to F was supported by post-experimental discussions with the participants regarding their strategies and how they remembered what each colour meant.

Figure 1. Schematic representation of the study phase presentation of ''Cleaning a Fish Tank'' (PsycheTruth, 2009) with a complementary timeline denoting the start and end time of each segment. Each video frame corresponds to one bar within the time line. The light-bordered frames represent R segments and the dark-bordered frames represent F segments. R and F instructions were randomly assigned to each segment on a participant-by-participant 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.

Test phase. Following the study phase, participants were tested using cued recall questions. Each question was presented in the centre of the computer screen immediately below the title of the tested video, which served as the retrieval cue. All questions related to a given video were presented prior to moving on to the next video, although the order in which individual questions were presented and the order in which the videos were tested were otherwise randomised. Participants were instructed to answer each question to the best of their ability using one or two words; although they were told they could use full sentences if they desired, no one did. To avoid participants dwelling for too long on a given question, they were instructed to respond *dnk* (*do not know*) or *idk* (*I don't know*) if faced with a question for which they could not even guess, although the use of this response was discouraged. Responses were visible on-screen immediately below the question and could be modified until the participant submitted their response by depressing RETURN on the keyboard, at which point the next question appeared.

The first author scored each response using an answer key created prior to data collection. During scoring the R or F instruction associated with the question was obscured to prevent bias. Because some questions had multiple acceptable answers that were unforeseeable prior to data collection, responses that the first author felt were correct—but were not listed on the original answer key—were flagged for independent review by an undergraduate research assistant and either added to the answer key or rejected. For example, the correct answer to the question *with what did the man flatten the chocolate chip cookies after rolling them?* was *his fingers*, but *his hands* was also deemed acceptable following independent review. Once all data had been collected, each question was rescored using *only* the answer key (that is, not adding any new responses) to ensure consistency. Misspellings were accepted as correct only if they were unambiguous. For example, *fingur* instead of *finger* would be deemed acceptable whereas *air* instead of *hair* would not be acceptable.

Results

The percentage of correct responses 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 and included in the overall count of the number of questions answered; excluding these responses as neither correct nor incorrect did not change the general pattern of results. This percentage of correct responses was then analysed as a function of instruction (R, F) using a one-way repeated-measures ANOVA. The analysis was significant, $F(1, 29) = 5.63$, $MSe = 83.52$, $p < .03$, revealing better performance for R segments ($M = 47.13\%$, $SE = 2.16\%$) than F segments ($M = 41.53\%$, $SE = 2.00\%$). The presence of a significant 5.60% R–F difference supports the viability of using continuous videos to study directed forgetting.

Discussion

Experiment 1 used a novel event method directed forgetting paradigm to investigate the ability to intentionally forget segments of a continuous visual event. Results supported the presence of a directed forgetting effect: Participants responded more accurately when tested for R segments than for F segments, as revealed by a 5.60% R–F difference in cued recall accuracy. This finding supports the assertion that intentional forgetting can occur for segments of continuous events, and is not restricted to discrete items or lists of words or pictures.

EXPERIMENT 2

Experiment 2 replicated the methods of Experiment 1 with the exception that memory was tested using true/false statements instead of cued recall questions. The use of true/false statements allowed us to separate the ability to discriminate between true/false statements from the response criterion employed during that discrimination process.

Method

Participants

A total of 30 undergraduate students (25 female) enrolled at the University of Arizona participated in this experiment for course credit. The majority of participants were right-handed (26 right, 4 left); their ages ranged from 18 to 21 years of age, with a mean of 18.47 years.

Stimuli and apparatus

The stimuli and apparatus were identical to those used in Experiment 1 with the exception that true/false statements were used instead of cued recall questions. An equal number of true and false statements were created for each segment: For two of the videos (Baking Assorted Cookies, Making Chocolate Pudding) four true and four false statements were created per segment and for the remaining two videos (Cleaning a Fish Tank, Getting Ready for Work) three true and three false statements were created per segment. The 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. The false statements were 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*). Overall, 224 test statements (112 true and 112 false) were created.

Procedure

Practice phase. The practice phase was identical to Experiment 1.

Study phase. The study phase was identical to Experiment 1.

Test phase. Following the study phase participants were presented with a series of true/false statements, one at a time, in the centre of the computer screen, each pertaining to a specific segment within the presented videos. Participants were tested for the content of one video at a time, although the statements for that video were presented in a random order and the videos were tested in a random order. To facilitate performance 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.

Results

The mean hits and false alarms are provided in Table 1. Using the procedure described by Donaldson (1992), non-parametric measures of

TABLE 1
Mean percentage of "true" responses as a function of instruction (remember, forget) and statement validity (true, false) for Experiments 2–4

	Statement validity			
	True		*False*	
Instruction	*M*	*SE*	*M*	*SE*
Experiment 2				
R	71	1	32	2
F	60	2	35	2
Experiment 3				
R	68	2	34	2
F	62	2	36	2
Experiment 4				
R	70	1	33	2
F	66	1	35	2

sensitivity (A') and response bias (B''_D) were calculated and analysed as a function of instruction (R, F) using paired t-tests. Values of A' range from chance ($A'=0.5$) to perfect performance ($A'=1.0$); values of B''_D range from liberal (requiring less "signal" to classify a test statement as true; $B''_D=-1.0$) to conservative (requiring more "signal" to classify a test statement as true; $B''_D=1.0$). As depicted in Figure 2, participants exhibited greater sensitivity, $t(29)=5.59$, $p<.01$, and responded more liberally, $t(29)=3.42$, $p<.01$, to statements about R segments than F segments. Analysis of the raw hits and false alarms instead of A' reveals a comparable pattern here as well as in the analyses reported for subsequent experiments.

Discussion

Experiment 2 replicated the findings of Experiment 1 using true/false statements instead of cued recall questions. Results supported the presence of a directed forgetting effect: Participants were more sensitive to statements about R segments than F segments. Interestingly, participants also responded more cautiously to statements about F segments than to statements about R segments, possibly because they were less able to retrieve additional, supporting details related to those segments.

To the extent that true/false statements may be an analogue of recognition memory, our findings could result in the speculation that the current directed forgetting effect is related to encoding differences for R and F segments at study (for a potential retrieval-based account, see Sahakyan, Waldum, Benjamin, & Bickett, 2009). Even so, we are disinclined to make a link between this finding and the item method paradigm that likewise tends to support differences in the encoding of R and F items (e.g., Basden, Basden, & Gargano, 1993). Whereas the memory instruction presented in an item method paradigm refers specifically to the item that is later tested, in our paradigm memory instructions were applied to an ongoing sequence of events that comprised a segment of a continuous video. Although these sequences depicted the individual details that were later tested, the memory instruction was never explicitly linked to these individual details but to the sequence as a whole. The R or F segments represented the aggregation of these individual details and in this manner were more similar in nature to how list method directed forgetting instructions apply to subgroups of studied materials as opposed to the individual elements themselves. Thus our paradigm shares features with both the item and list method paradigms and yet is different from both.

One concern regarding our findings might be that participants were able to guess which statements were true and which were false on the basis of previously acquired knowledge, or schemas, concerning the tested event. Although we can counter this concern by noting that responses

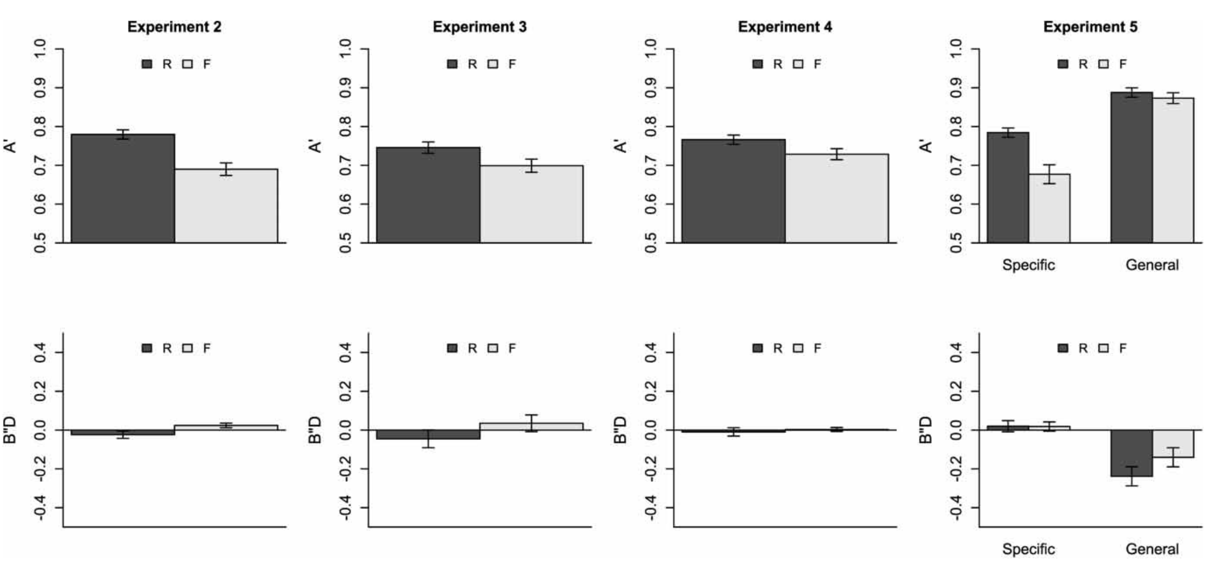

Figure 2. Sensitivity (A') and Response Bias (B''_D) as a function of instruction (R, F) and specificity (specific, general) as applicable for Experiments 2, 3, 4, and 5; error bars represent one standard error of the mean.

based on activated schemas would not be expected to produce differences in accuracy between R and F segments, we nevertheless decided to address this issue directly. To get a baseline measurement of the discriminability between the true and false statements used in this experiment, we ran six new participants in a replication of Experiment 1 that eliminated the practice and study phases. These participants were presented only with the test phase, during which they were asked to guess whether each statement was true or false on the basis of the provided title and *without* having watched the videos. *A'* and B''_D were calculated as described above. Performance was very close to chance ($M = 0.55$, $SE = 0.04$) and was well below the experimental group (regardless of memory instruction), although participants did demonstrate a slightly liberal response bias ($M = -0.10$, $SE = 0.06$). Clearly the performance observed in Experiment 1 cannot be accounted for in any large part by prior, extra-experimental, knowledge of events depicted in the videos.

We interpret our findings as evidence that participants can selectively exclude to-be-forgotten segments from the encoding of an otherwise continuous sequence of visual events. A more mundane explanation that must first be ruled out, however, is the possibility that participants used the coloured border that served as the F instruction as a signal to redirect visual attention away from the video. This would account for poorer memory performance for F segments than R segments, but not for the reason we have presumed. Based on the results of Experiment 2 we think it unlikely that participants simply failed to observe the to-be-forgotten segments. If this had been the case, performance for F segments would have approximated performance in the no-video baseline group described above (wherein participants responded to statements regarding segments they had not observed). This was not the case. Nevertheless, to further rule out this possibility, Experiment 3 replicated the basic paradigm developed in Experiments 1 and 2 except that participants were required to attend the videos at all times to respond to visual targets that appeared unpredictably during both R and F segments.

EXPERIMENT 3

In Experiment 3 small triangles served as visual targets that required a speeded discrimination response to indicate the location of their apex. These were superimposed on the video during the study phase. To ensure temporal unpredictability, these targets were presented at intervals ranging from 1 to 32 seconds relative to segment onset. Due to the difficulty associated with detecting and discriminating a small visual target within a complex scene, this task encouraged visual attention to remain within the viewing area at all times, even during F segments. To encourage attention to remain roughly centralised, half of the targets appeared at centre, with the remaining targets equally distributed between the near and distant periphery of the video.

Method

Participants

A total of 30 undergraduate students (20 female) enrolled at the University of Arizona participated in this experiment for course credit. The majority of participants were right-handed (26 right, 4 left); their ages ranged from 18 to 22 years of age, with a mean of 19.13 years.

Stimuli and apparatus

The stimuli and apparatus were identical to those used in Experiment 2 with the exception that Experiment 3 included a blue (RGB values 0,100,0) isosceles triangle that served as a visual target requiring a speeded button press response. This triangle measured 30 pixels along its base and 30 pixels from base to tip presented on its side such that it acted as an arrowhead, with the apex pointing to the left or right of the computer screen.

Procedure

Practice phase. The practice phase was identical to Experiment 2, with the exception that it incorporated the discrimination task used during the study phase (see below).

Study phase. The study phase replicated the procedures of Experiment 2 with the exception that a small blue triangle with its apex pointing to the left or to the right was occasionally superimposed on each video for approximately 600 ms (15 frames). Participants were required to indicate the direction in which these triangles pointed by depressing the "f" key if the triangle pointed to the left and the "j" key if the triangle pointed to the right, regardless of where the triangle appeared relative to centre. Participants rested the appropriate index finger on each button at all times so as to be prepared to respond.

Three triangles were presented per segment. To maximise the unpredictability of the target and to avoid targets appearing too rapidly following each other, one of these triangles was presented after a short delay, one after an intermediate delay, and one after a long delay; each delay category included two possible intervals relative to segment onset that were sampled randomly but with equal probability: short (1 second, 2 seconds), intermediate (4 seconds, 8 seconds), and long (16 seconds, 32 seconds). Thus each video segment contained a single target from each delay category (e.g., a target could be presented 1 second or 2 seconds following the onset of a given segment, but not both), with the exact timing staggered due to the sampling of two possible intervals within each delay category. The video viewing area was conceptually separated into a grid comprising 25 squares measuring 30 pixels $\times$ 30 pixels each. This grid was further separated into three regions: centre (the location immediately at centre), middle (the ring of 8 squares surrounding the centre position), and outer (the outer ring of 16 squares). Targets were centred in a random square located within the selected region. To encourage attention to remain generally centralised, targets were most likely to appear in the centre region (50% of all targets) with the remaining targets equally distributed between the middle and outer regions (25% of all targets, each). A total of 96 targets were presented in this manner.

Collapsing across target direction, the study phase was conceptualised as a 2 (instruction: R, F) $\times$ 6 (interval: 1, 2, 4, 8, 16, 32 seconds) $\times$ 3 (region: centre, middle, outer) within-participants design. For the analyses, the factor of interval was collapsed from six to three levels (short, intermediate, and long). These factors were balanced across (not within) the videos; each video contained four R and four F segments and eight short, eight intermediate, and eight long targets.

Test phase. The test phase was identical to that described for Experiment 2.

Results

Discrimination reaction times (RTs)

Even though the discrimination targets were included primarily to ensure that participants attended to the computer screen at all times—even when instructed to forget the current segment—we nevertheless analysed these RTs as a function of instruction (R, F), interval (short, intermediate, long), and region (centre, middle, outer) using a three-way repeated-measures ANOVA; these data are presented in Table 2. Analyses conducted using log- or inverse-transformed RTs produced a pattern comparable to the analyses reported below. The main effect of instruction, $F(1, 29) = 10.05$, $MSe = 33195.29$, $p < .01$, revealed slower responses to target arrows presented during R segments than F segments. The main effects of both interval, $F(2, 58) = 15.29$, $MSe = 16093.59$, $p < .01$, and region, $F(2, 58) = 10.82$, $MSe = 15451.13$, $p < .01$, were also significant representing a tendency for participants to respond most rapidly to targets presented at the intermediate interval and for RTs to increase with increasing visual eccentricity. None of the interactions were significant, all ps $> .09$.

Target discrimination accuracy

Because an RT analysis is meaningless without knowing whether there is evidence of a speed–accuracy trade-off, an identical analysis was conducted on the percent correct target discrimination responses (also presented in Table 2). The only effect to reach significance was location, $F(2, 58) = 14.45$, $MSe = 0.03$, $p < .01$, demonstrating better accuracy for discriminating the direction of the triangle apex for targets presented at centre (the most likely target location) than elsewhere. This finding contradicts a speed–accuracy trade-off, given that targets presented at centre were also responded to the fastest. No other effects or interactions even approached significance (all Fs < 1).

Signal detection analysis

The mean hits and false alarms are provided in Table 1. As in Experiment 2, non-parametric measures of sensitivity (A') and response bias (B''_D) were calculated and analysed as a function of instruction (R, F) using paired t-tests (see

TABLE 2

Mean discrimination RTs (ms) and accuracies (%) in Experiment 3 as a function of region (centre, middle, outer), instruction (R, F), and interval (short, intermediate, long)

	RT (ms)						*Accuracy (%)*					
	Short		*Intermediate*		*Long*		*Short*		*Intermediate*		*Long*	
Instruction	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
Centre												
R	852	37	738	23	803	36	91	3	91	2	92	2
F	787	25	721	22	793	34	88	3	93	1	92	2
Middle												
R	839	32	769	31	848	37	90	3	88	3	84	4
F	769	21	727	26	814	37	89	3	88	4	86	3
Outer												
R	916	42	849	41	850	36	84	4	86	3	79	3
F	854	37	766	31	817	31	82	4	84	4	82	4

Figure 2). Once again, participants exhibited greater sensitivity, $t(29) = 3.46$, $p < .01$, and responded more liberally, $t(29) = 2.06$, $p < .05$, to statements about R segments than F segments.

Discussion

Experiment 3 replicated the results of Experiment 2 using a speeded visual target discrimination task that required participants to maintain their visual attention on the video at all times—even when instructed to forget a given segment. Participants were fastest and most accurate to respond to visual targets that were presented at the expected (most frequent) target location (centre), and this tendency was not changed as a function of memory instruction. These results suggest that participants did, in fact, maintain their attention at centre throughout the presentation of both R and F segments. Even with visual attention focused on the video throughout all segments, the pattern of memory performance remained essentially unchanged from Experiment 2: Participants showed greater sensitivity when tested for R segments than F segments.

Interestingly, in the target discrimination data there was a main effect of memory instruction on RT, such that participants were faster to respond to targets presented during F than R segments. The fact that memory instruction did not interact with target location means that there is no evidence that participants adopted a "looking away" strategy during F segments. Nevertheless, the finding of overall slower RTs during F than R segments supports the hypothesis that the directed forgetting effect in memory was likely related to differential encoding of these segments. Where RT to respond to the target can be used to index the cognitive demands associated with remembering and forgetting (Fawcett & Taylor, 2008; see Kahneman, 1973), our pattern of target RTs suggests that encoding the R segments was more demanding than encoding the F segments. This finding is in apparent contrast to results obtained using a simple detection task embedded in a more typical *item method* directed forgetting paradigm. Fawcett and Taylor (2008) had participants respond to a visual detection probe ("*") presented 1400 ms, 1800 ms, or 2600 ms following each study phase memory instruction in a task that presented discrete trials that contained a single word each. Participants were *slower* to respond to probes following F than R instructions at the 1400 ms and

1800 ms intervals. Fawcett and Taylor (2008) suggested that instantiating an F instruction in an item method directed forgetting task is *more* cognitively demanding than instantiating an R instruction—an assertion that converges with neuroimaging results that implicate frontal lobe involvement following F instructions (e.g., Wylie, Foxe, & Taylor, 2008).

There are two primary reasons why the current finding of longer target discrimination RTs during R than F segments need not conflict with the apparently opposite RT results reported by Fawcett and Taylor (2008). First, Fawcett and Taylor (2008) measured responses immediately following each memory instruction at delays far shorter than most used in the current investigation. Indeed, they argued that the purpose of the active mechanism engaged by a F instruction (while briefly cognitively demanding) is ultimately to free processing resources for other tasks—including the rehearsal of previous R items, responding to a secondary task, or performing some other diversionary cognitive process (see Fawcett & Taylor, 2012; Sahakyan & Kelley, 2002). The finding of faster discrimination RTs during F than R segments could be interpreted as support for the view that at a relatively long interval following the F instruction, processing resources have been successfully diverted from the to-be-forgotten content of the video and to the discrimination task instead. Second, Fawcett and Taylor (2008) used discrete trials that instructed participants to remember or forget information that was already actively represented in working memory in anticipation of the impending memory instruction. The notion was that once the study word had been presented, participants would maintain that word in working memory during the brief period preceding the memory instruction; following an R instruction the study word would be rehearsed whereas following an F instruction the active mechanisms described above would be engaged to stop the continued rehearsal of the item and thereby interfere with its successful encoding to long-term memory (Fawcett & Taylor, 2009). Therefore, according to Fawcett and Taylor (2008, 2012; Taylor 2005) the act of forgetting in a typical item method paradigm largely requires the participant to exert control over the *current contents* of working memory. In contrast, in the current experiment each R and F memory instruction was concurrent with the studied information. Instantiating an F instruction in this instance requires the participant to control *access to* working memory.

EXPERIMENT 4

In Experiment 3 participants were encouraged to attend each video segment by virtue of requiring a speeded discrimination response to targets that appeared briefly during R and F segments. The results of that experiment support the notion that participants did, in fact, maintain attention on the video segments on both R and F trials so that differences in attentional locus cannot account for the directed forgetting effect in memory. In Experiment 4 we extended this investigation by requiring participants to conceptually separate each video into subjectively determined subordinate events by depressing the spacebar whenever they conceptualised an action as representing the beginning of a new event (see Zacks & Tversky, 2001). For example, while watching the practice video (Folding Laundry), one might depress the spacebar each time a new article of clothing was removed from the laundry basket—or even each time an individual fold was made to a given article of clothing—depending on the specificity with which one has chosen to define what constitutes an event. While determining event boundaries is necessarily subjective, it requires that participants attend the video at all times, and that they process the content of the video *conceptually* to determine the appropriate time to make each event-segmentation response. Because this task was conducted throughout both R and F segments, this ensured that all portions of the video received some degree of conceptual encoding, especially to the extent that a similar number of event boundaries are assigned to R and F segments.

Method

Participants

A total of 31 undergraduate students (24 female) enrolled at the University of Arizona participated in this experiment for course credit. The majority of participants were right-handed (24 right, 7 left); their ages ranged from 18 to 23 years of age, with a mean of 18.83 years.

Stimuli and apparatus

The stimuli and apparatus were identical to those used in Experiment 2.

Procedure

Practice phase. The practice phase was identical to that described for Experiment 2, with the exception that participants engaged in an event-segmentation task while watching the practice video.

Study phase. The study phase was identical to that described for Experiment 2, with the exception that participants were instructed to keep the index finger of their dominant hand on the spacebar at all times and to depress it whenever they determined that a new event or action had begun. In accordance with prior work investigating event segmentation (see Zacks & Tversky, 2001) participants were explicitly instructed that there were no right or wrong times at which to depress the spacebar and were provided with a few examples demonstrating how a single event might be segmented in multiple ways depending on the criterion employed and how the terms were defined (e.g., *While folding laundry, the act of folding a shirt could be a single action or each individual fold could be a separate action*).

Recognition phase. The recognition phase was identical to that described for Experiment 2.

Results

Event-segmentation analysis

Although included primarily to ensure that participants conceptually encoded the videos at all times—even when instructed to forget the current segment—the average number of event segmentation responses was analysed as a function of instruction (R, F) using a one-way repeated-measures ANOVA. The effect of memory instruction was not significant, $F(1, 30) = 1.06$, $MSe = 0.45$, $p > .31$, with an equivalent mean number of event boundaries during F segments ($M = 4.85$, $SE = 0.44$) and R segments ($M = 4.68$, $SE = 0.45$). The fact that participants placed event boundaries at a similar rate during F segments and R segments supports the notion that participants encoded these segments at a similar conceptual level. This interpretation is supported by Figure 3, which depicts the percentage of segmentation responses made across participants for each video as a function of time and instruction. Spikes indicate points at which participants tended to agree that an event boundary had occurred. Importantly, the lines depicting R and F trials are highly similar, overlapping almost entirely at points.

Signal detection analysis

The mean hits and false alarms are provided in Table 1. As in the previous experiments, non-parametric measures of sensitivity (A') and response bias (B''_D) were calculated and analysed as a function of instruction (R, F) using paired t-tests (see Figure 2). Although participants remained more sensitive to statements about R segments than F segments, $t(30) = 2.74$, $p < .02$, they did not differ in their response bias, $t(30) = 0.62$, $p > .54$.

Discussion

Memory performance in Experiment 4 replicated the findings of Experiments 2 and 3 even while participants engaged in an event-segmentation task designed to ensure that each video was encoded conceptually during both R and F segments. Analysis of the secondary event-segmentation task revealed that participants placed an equivalent number of event boundaries in R and F segments—a task that requires careful and attentive analysis of the events occurring in each video. Despite requiring participants to process F segments at a conceptual level, they still responded with greater sensitivity when tested for R segments than F segments. Clearly, the directed forgetting effect for continuous event sequences cannot be easily dismissed as an artefact of a "looking away" strategy on F trials.

Thus far the directed forgetting effect has been discussed as the difference in memory performance observed for R and F video segments. However, the directed forgetting effect is believed to arise from an aggregation of processes acting to strengthen memory for R items as well as processes acting to weaken memory for F items. For this reason, the R–F difference score that defines a directed forgetting effect can be conceptualised as consisting of *costs* (worse memory for F items) plus *benefits* (better memory for

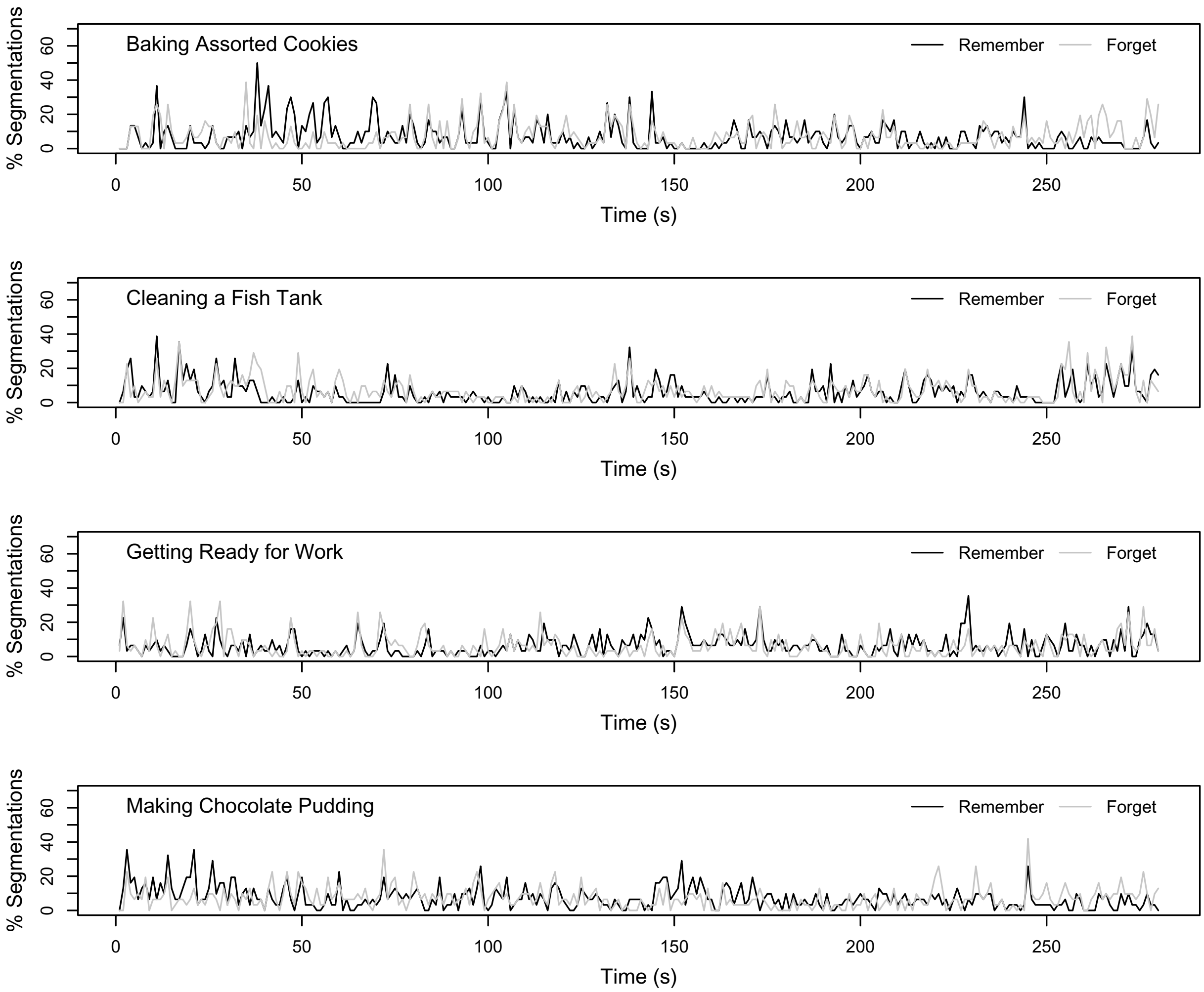

Figure 3. The percentage of segmentation responses made for each video in Experiment 4 as a function of time in seconds and instruction (R, F); data are presented with a granularity of 1 second.

R items). To separate costs from benefits, in a between-participants manipulation, a remember-all control group is used to measure memory performance when all items must be committed to memory. Benefits are measured as better performance for R items in the directed forgetting task compared to the remember-all baseline, and costs are reflected in worse performance for F items in the directed forgetting task compared to the remember-all baseline (e.g., Basden & Basden, 1996; Sahakyan & Foster, 2009).

An additional 17 participants were run in a complete replication of Experiment 4 with the exception that the random changes in the coloured border from green to purple and vice versa was not ascribed any meaning. Participants were instructed to remember every segment of the videos that they watched, while also engaging in the event segmentation task. A signal detection analysis was performed on their test data resulting in a mean *A'* score of 0.77 ($SE = 0.01$). Planned *t*-tests evaluated whether *A'* scores for R or F segments in the experimental version of the task differed from the overall *A'* scores for the control version of the task. These analyses revealed significant *costs*, with worse discriminability of F segments than the remember-all control segments, $t(46) = 2.12$, $p < .04$. There was no evidence of *benefits*, such that the discriminability of R segments was statistically indistinguishable from that of the remember-all control segments, $t(46) = 0.35$, $p > .73$. This suggests that our overall directed forgetting effect, reflected in the difference between R and F performance, reflects primarily costs to memory associated with the instruction to forget.[2]

The fact that we observed costs without benefits may appear at odds with the results of Sahakyan and Foster (2009) who observed both *costs* and *benefits* for list- and item method directed forgetting tasks. It is possible that the absence of benefits is a defining feature of event method directed forgetting. More likely, however, is that benefits depend on the dependent variable that is used to assess memory performance. Both Basden and Basden (1996) and Sahakyan and Foster (2009) used a recall task to measure memory performance in their experiments and observed both costs and benefits. In contrast, across three experiments Taylor and Fawcett (2012) observed costs without benefits in the context of an item method directed forgetting task that measured yes-no recognition performance. To the extent that true/false statements are more akin to recognition than recall, it follows that the benefits of directed forgetting may be larger and more robust for tasks that depend on recall rather than recognition. In any case it is clear that the costs of the forget instruction are robust across paradigms and dependent measures, supporting our contention that the difference in performance for R and F segments is due primarily to intentional forgetting of the F segments.

To compare the directed forgetting effect across experiments, the *A'* data were collapsed and analysed using a mixed-effects ANOVA with experiment (Experiment 2, Experiment 3, Experiment 4) as the between-participants factor and instruction (R, F) as the within-participants factor. Although the main effect of experiment was not significant, $F(2, 88) = 1.02$, $MSe = 0.01$, $p > .36$, the main effect of instruction, $F(1, 88) = 48.00$, $MSe < 0.01$, $p < .01$, and the interaction of experiment and instruction were significant, $F(2, 88) = 3.75$, $MSe < 0.01$, $p < .03$. The magnitude of the directed forgetting effect was larger in Experiment 2 than in Experiment 3, $t(58) = 2.06$, $p < .05$, or Experiment 4, $t(59) = 2.49$, $p < .02$; the magnitude of the directed forgetting effect in Experiments 3 and 4 did not differ, $t(59) = 0.48$, $p > .63$.

Figure 2 suggests that the difference between Experiment 2 and Experiments 3 and 4 may arise from separate sources. Participants in Experiment 3 demonstrated *diminished* sensitivity to statements testing R segments ($A' = 0.75$) relative to participants in Experiment 2 ($A' = 0.78$) whereas sensitivity to statements testing F segments did not differ between these experiments ($A' = 0.69$ for both). This suggests that the Experiment 3 speeded target discrimination task might have interfered with the retention of R segments by disrupting the participant's encoding or rehearsal strategy. Part of this rehearsal strategy might involve drawing inferences. For example, viewers often generate predictive inferences while observing an event or listening to a discourse (e.g., Magliano, Dijkstra, & Zwaan, 1996). These inferences facilitate comprehension and therefore retention of those materials. We speculate that it is possible that the cognitive demands of the speeded target discrimination task impaired rehearsal (including inference generation) during R segments contributing to the relative reduction in sensitivity between Experiments 2 and 3. In contrast, participants in Experiment 4 demonstrated *greater* sensitivity to statements testing F segments ($A' = 0.73$) than participants in Experiment 2 ($A' = 0.69$) whereas sensitivity to statements testing R segments differed minimally between these experiments ($A' = 0.78$ vs $A' = 0.77$). This suggests that the Experiment 4 event-segmentation task improved memory for F segments by forcing participants to adopt a conceptual encoding strategy. The fact that the event-segmentation task (in Experiment 4) did not negate the directed forgetting effect completely may represent the contribution of an active rehearsal strategy that favoured R over F segments.

EXPERIMENT 5

Having demonstrated that memory for relatively specific details of a visual event are impaired for F segments relative to R segments, it is reasonable to consider whether the same is true for relatively general details. At least with respect to *unintentional* forgetting, specific and general details appear to be forgotten at different rates. For

[2] Our analysis of the costs and benefits in the current paradigm have focused on *A'* because this is the metric for which the comparisons against a remember-all baseline condition are most theoretically meaningful; however, we also calculated and analysed the B''_D scores. Performance in the control condition was relatively unbiased, with a mean B''_D score of -0.01 ($SE = 0.02$). Comparisons revealed that the response bias observed in the control condition did not differ from either the response bias observed for R segments, $t(46) = 0.07$, $p > .94$, or for F segments, $t(46) = 0.78$, $p > .43$.

example, Dorfman and Mandler (1994) found that whereas specific item information about a categorised word list was lost relatively quickly as the delay between study and test increased from no delay to a week, gist memory for the categories remained relatively stable. The finding that general information is more resilient to unintentional forgetting may imply that it is similarly resilient to intentional forgetting. By exploring how intentional forgetting influences details relative to general information we may also learn more about the granularity of the control processes involved. Does cognitive control over memory allow us to selectively forget the details of an unwanted event or do we forget the whole event?

Experiment 5 replicated the methods developed in Experiment 2 with the exception that half of the true/false statements were modified such that they tested relatively specific details (*the women added two measuring cups of milk to the mixture*) whereas the remaining half tested relatively general details (*the women added milk to the mixture*). We hypothesised that participants would derive a general gist representation for each event that was presented, even those contained in segments they had been instructed to forget. According to this hypothesis we predicted a pattern of results similar in direction and magnitude to Experiment 2 for the specific statements and a pattern of results similar in direction albeit smaller (or non-significant) for the general statements. Because we were primarily concerned with the magnitude of the directed forgetting effect as a function of granularity, there was no need to include a secondary task during the study phase. Even though our previous experiments ruled out any major influence of a "looking away" strategy on the directed forgetting effect, to the extent that such a strategy emerged on F trials due to the lack of a secondary task, memory for both gist and specific details would be affected.

Method

Participants

A total of 20 undergraduate students (11 female) enrolled at the University of Arizona participated in this experiment for course credit. Participants were entirely right-handed and ranged from 18 to 24 years of age, with a mean of 18.95 years.

Stimuli and apparatus

The stimuli and apparatus were identical to Experiment 2 with the exception that the true/false statements were modified such that half tested relatively specific details (*the women added two measuring cups of milk to the mixture*) whereas the remaining half tested relatively general details (*the women added milk to the mixture*). In most cases specific statements were taken directly from Experiments 2, 3, and 4 with general questions created by removing the specific details from these statements. Each video was tested using a total of 8 test statements per segment (2 true specific, 2 true general, 2 false general, 2 false specific) for a total of 64 test statements per video and 256 test statements overall.

Procedure

Practice phase. The practice phase was identical to Experiment 2.

Study phase. The study phase was identical to Experiment 2.

Test phase. The test phase used true/false statements and was identical to Experiment 2 with the exception noted above that half of these statements tested more specific details whereas the remaining half tested more general details. General and specific test statements were presented randomly.

Results

The mean hits and false alarms are provided in Table 3. Non-parametric measures of sensitivity (A') and response bias (B''_D) were calculated and analysed as a function of instruction (R, F) and relative specificity (specific, general) using separate two-way repeated-measures ANOVAs (see Figure 2). For the A' analysis there was a main effect of instruction, $F(1, 19) = 15.66$, $MSe = 0.01$, $p < .01$, with greater sensitivity to statements about R segments than F segments. The main effect of relative specificity, $F(1, 19) = 100.07$, $MSe = 0.01$, $p < .01$, was also significant, with greater sensitivity to general statements than specific statements. Importantly, these effects were qualified by a significant instruction × relative specificity interaction, $F(1, 19) = 9.84$, $MSe = 0.01$, $p < .01$. Planned contrasts on A'

TABLE 3
Mean percentage of "true" responses as a function of instruction (remember, forget), specificity (specific, general) and statement validity (true, false) for Experiment 5

	Statement validity			
	True		*False*	
Instruction	*M*	*SE*	*M*	*SE*
Specific				
R	70	2	29	2
F	59	3	35	2
General				
R	88	1	23	2
F	85	2	23	2

revealed a significant 0.11 directed forgetting effect for specific statements, $t(19) = 3.96$, $p < .01$. With a mean R–F difference of 0.02, there was no significant directed forgetting effect for general statements, $t(19) = 1.08$, $p > .29$.

The B''_D analysis revealed only a significant main effect of relative specificity, $F(1, 19) = 59.76$, $MSe = 0.05$, $p < .01$, with participants employing a more liberal response bias for general statements than specific statements; neither the main effect of instruction, $F(1, 19) = 2.01$, $MSe = 0.11$, $p > .17$, nor the instruction × specificity interaction, $F(1, 19) = 0.57$, $MSe = 0.02$, $p > .45$, was significant.

Discussion

Experiment 5 explored the level of specificity at which the directed forgetting effect could be measured. The *A'* difference for the specific test statements (*A'* = 0.11) approximated the difference observed in Experiment 2 (*A'* = 0.09). However, the R–F difference was numerically smaller (*A'* = 0.02) for general test statements and failed to reach significance. This finding is interesting because it suggests that successful intentional forgetting is not an all-or-nothing phenomenon. General information is more resilient to intentional forgetting than relatively specific information. While important, this finding is not shocking. There is a rich literature exploring instructions to disregard inadmissible information in court and a comparable literature dealing with instructions to disregard information in social settings that have periodically found that despite earnest attempts at suppressing such information, the to-be-forgotten information continues to influence behaviour (see Isbell, Smith, & Wyer, 1998; Johnson, 1998; Kassin & Studebaker, 1998). The current findings suggest that the locus of this influence could reside in a residual gist-based trace.

As was done for Experiment 4, a remember-all control study was also conducted for Experiment 5. A total of 12 participants completed an exact replication of Experiment 5, except that no meaning was ascribed to the coloured border; participants were instructed to commit all segments to memory. The mean *A'* score was 0.77 ($SE = 0.01$) for specific test statements and 0.90 ($SE = 0.01$) for general test statements. Planned *t*-tests on the *A'* scores for the specific test statements revealed significant *costs*, $t(30) = 2.92$, $p < .01$, but no evidence of *benefits*, $t(30) = 0.54$, $p > .59$. Given that the general test statements showed no significant overall directed forgetting effect (i.e., costs + benefits ≈ 0), it is not surprising that when considered separately there were neither significant *costs*, $t(30) = 1.12$, $p > .27$, nor *benefits*, $t(30) = 0.47$, $p > .64$ for these statements.[3]

To rule out any effects of pre-existing schematic representations of the event itself we ran 10 new participants in a replication of Experiment 5 that presented only the test statements, without any prior exposure to the videos. Performance for the specific condition was numerically equivalent to chance ($M = 0.50$, $SE = 0.02$), whereas performance for the general condition was only slightly above chance ($M = 0.59$, $SE = 0.04$). Liberal response biases were observed for both the specific ($M = -0.16$, $SE = 0.12$) and general ($M = -0.51$, $SE = 0.07$) conditions. Although participants were slightly better at guessing the correct responses for the general test statements, performance was still far below the performance in both the R (.90) and F (.88) conditions of the experimental group. In addition to removing serious concerns regarding the contribution of guessing to test performance, these findings also demonstrate that participants were attending to the videos: Had they looked away during F

[3] As in Experiment 4, our analysis of the costs and benefits in the current paradigm have focused on *A'*; however, we also calculated and analysed the B''_D scores. Performance in the control condition was relatively unbiased for the specific test statements ($M = 0.02$, $SE = 0.03$) and relatively liberal for the general test statements ($M = -0.24$, $SE = 0.05$). Importantly, comparisons revealed response bias for both the specific and general test statements to be statistically comparable to the values calculated for the R and F segments (for the general baseline vs general forget comparison, $t(30) = 1.79$, $p > .08$; all other *p*s > .61).

segments, performance for those segments would have approximated the no-video control group.

Our finding of a significant directed forgetting effect for specific but not for general information may appear at odds with Joslyn and Oakes (2005) who observed a significant directed forgetting effect in their diary study for general information but *not* for specific details. Joslyn and Oakes' (2005) measurement of gist memory comprised a two-word label (e.g., Shopping Trip) provided by participants for each event they recorded in their diary. Given the broad nature of the events studied in their experiment, it is probable that their measurement of gist rested at a more general level of analysis than the questions in the general condition of the current experiment, which typically tested memory for subordinate events. To put this into context, the titles used to describe each video (e.g., Baking Cookies) would constitute a gist statement equivalent to that used in Joslyn and Oakes' (2005) diary experiment. Moreover, Joslyn and Oakes (2005) tested memory for specific details only when the gist statement for the relevant event (e.g., Shopping Trip) had already been recalled; due to the fact that the specific information defined the event (and helped supply the event title used to test gist) it is not surprising that—should the gist of the event be recalled—central features of the event (e.g., time of day) would also be recalled. Our general statements were at a finer grain of analysis than used by Joslyn and Oakes (2005) and we tested specific statements whether or not the gist statements were answered correctly.

GENERAL DISCUSSION

Five experiments investigated the influence of intentional forgetting on subsequent memory for continuous visual events. In each experiment four videos were presented surrounded by a coloured border that changed from green to purple at variable intervals: Participants were instructed that whenever the border was green they would need to remember everything that was presented for a later test (R segment); whenever the border was purple they could forget everything that was presented (F segment). Following the presentation of the study videos, participants demonstrated better memory for information presented during R segments than information presented during F segments using cued recall questions (Experiment 1) and true/false statements (Experiments 2, 3, 4, and 5); this pattern was replicated even in the presence of tasks intended to maintain visual attention on the video (Experiment 3) or encourage conceptual processing of the video (Experiment 4). The R–F differences observed in these experiments were limited to relatively *specific* as opposed to relatively *general* test statements (Experiment 5).

These experiments extend the framework of intentional forgetting from static information such as pictures or words to dynamic information such as continuous visual events. This is a critical step if intentional forgetting is to be applied to real-world experiences. Equally important is the revelation that intentional forgetting does not reduce memory in an all-or-nothing fashion. Rather, there is a graded loss of information: largest for relatively specific details and smallest or even absent for relatively general details. In Experiments 4 and 5 these differences were entirely attributable to the *costs* without *benefits* of intentional forgetting (see also Taylor & Fawcett, 2012).

Interestingly, past research using semantically related word lists (e.g., Golding, Long & MacLeod, 1994) and structured narratives (e.g., Geiselman, 1974, 1977) has found that meaningful connections between to-be-remembered and to-be-forgotten information undermine intentional forgetting (see also MacLeod, 1998). This may arise because participants integrate F information with the R information to maintain cohesion. To the extent that the R and F segments that comprised our visual vignettes were likewise related through a common implied storyline, one might have reasonably expected no directed forgetting effect. And yet our effects were remarkably robust.

Although we favour a differential rehearsal interpretation of our directed forgetting effects, with better encoding of R segments than F segments, we do not exclude the possibility that limiting access of the F segments to working memory involves one or more active processes. When considering the current results one must remain mindful that remembering and forgetting in an intentional forgetting paradigm represent independent strategies/processes that combine to produce a directed forgetting effect. We have concluded that participants likely engaged in an active rehearsal strategy during R segments, accounting for slower responses to visual discrimination targets; however, the manner in which participants behaved during F segments remains uncertain. Recent research using the list method has begun exploring the contribution of diversionary thoughts to the intentional forgetting of lists (e.g., Sahakyan & Kelley, 2002; for a similar idea using the item method see Fawcett & Taylor, 2012). While it seems clear that participants did not adopt a ‘‘looking away’’ strategy during the presentation of F segments in the current study, it is possible that they attempted to think about something else as a means of ignoring the to-be-forgotten information. To the extent that reallocating attention from external to internal representations might be expected to slow RTs to discrimination targets presented during F compared to R segments, the use of a diversionary strategy does not seem likely. Such a strategy would also have difficulty explaining the survival of the directed forgetting effect in Experiment 4: Even if diversionary thoughts played some role in the participants' strategy, the act of segmenting each F segment still required conceptual consideration of the events for which they were tested. In contrast, it seems possible that the encoding (and/or rehearsal) of F segments could be actively controlled as suggested by some researchers using the item method (e.g., Fawcett & Taylor, 2010; Taylor, 2005; Ullsperger, Mecklinger, & Müller, 2000; Van Hooff, & Ford, 2011; Zacks et al., 1996).

Regardless of exactly how the directed forgetting is achieved in our event method paradigm, it is clear that intentional forgetting can occur even for segments of continuous visual events. The current data demonstrate that our paradigm is capable of producing a robust directed forgetting effect as measured by both cued recall and true/false statements. This directed forgetting effect is not attributable to a ‘‘looking away’’ strategy during F segments, as it occurs even during the performance of a secondary task that requires sustained visual or conceptual encoding of the F segments. Even so, this robust directed forgetting effect extends to relatively specific but not to relatively general information, arguing that the costs of forgetting may be due to changes in how specific information is encoded into memory; the gist trace may be relatively unaffected by an F instruction. We propose that specific information regarding the R segments (but not the F segments) is retained by limiting access of to-be-forgotten segments to working memory resources and also by selectively rehearsing to-be-remembered segments.

Manuscript received 3 April 2012
Manuscript accepted 6 November 2012
First published online 9 January 2013

REFERENCES

Barnier, A., Conway, M. A., Mayoh, L., Speyer, J., Avizmil, O., & Harris, C. B. (2007). Directed forgetting of recently recalled autobiographical memories. *Journal of Experimental Psychology: General*, *136*, 301–322.

Basden, B. H., & Basden, D. R. (1996). Directed forgetting: Further comparisons of the item and list methods. *Memory*, *4*, 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: Lawrence Erlbaum Associates Inc.

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.

Bjork, R. A. (1972). Theoretical implications of directed forgetting. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington, DC: Winston & Sons.

Bjork, R. A., LaBerge, D., & Legrand, R. (1968). The modification of short-term memory through instructions to forget. *Psychonomic Science*, *10*, 55–56.

Brown, J. (1954). The nature of set-to-learn of intra-material interference in immediate memory. *The Quarterly Journal of Experimental Psychology*, *6*, 141–148.

Burwitz, L. (1974). Proactive interference and directed forgetting in short-term motor memory. *Journal of Experimental Psychology*, *102*, 799–805.

Cicero, M. T. (2009). *De oratore*. New York, NY: General Books. [Original work published 1840.]

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

Dorfman, J., & Mandler, G. (1994). Implicit and explicit forgetting: When is gist remembered? *The Quarterly Journal of Experimental Psychology*, *47A*, 651–672.

Earles, J. L., & Kersten, A. W. (2002). Directed forgetting of actions by younger and older adults. *Psychonomic Bulletin & Review*, *9*, 383–388.

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

Fawcett, J. M., & Taylor, T. L. (2009, December). *Comparing item method directed forgetting to the active suppression of rehearsal: Further evidence from reaction time probes*. Poster presented at the Annual Meeting of the Psychonomic Society, Boston, MA.

Fawcett, J. M., & Taylor, T. L. (2010). Directed forgetting shares mechanisms with attentional withdrawal but not stop-signal inhibition. *Memory & Cognition*, *38*, 796–807.

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.

Geiselman, R. E. (1974). Positive forgetting of sentence material. *Memory & 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.
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.
Isbell, L. M., Smith, H., & Wyer, R. S. (1998). Consequences of attempts to disregard social information. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 289–320). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
James, W. (1950). *The principles of psychology* (Vol. 2). New York, NY: Dover. [Original work published 1890.]
Johnson, H. M. (1998). Disregarding information in text. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional Forgetting: Interdisciplinary Approaches* (pp. 219–237). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
Joslyn, S. L., & Oakes, M. A. (2005). Directed forgetting of autobiographical events. *Memory & Cognition*, *33*, 577–587.
Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.
Kassin, S. M., & Studebaker, C. A. (1998). Instructions to disregard and the jury: Curative and paradoxical effects. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional Forgetting: Interdisciplinary Approaches.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
MacLeod, C. M. (1998). Directed forgetting. In J. M. Golding & C. M. MacLeod (Eds.), *Intentional forgetting: Interdisciplinary approaches* (pp. 1–57). Mahwah, NJ: Lawrence Erlbaum Associates Inc.
Magliano, J. P., Dijkstra, K., & Zwaan, R. A. (1996). Generating predictive inferences while viewing a movie. *Discourse Processes*, *22*, 199–224.
Muther, W. S. (1965). Erasure or partitioning in short-term memory. *Psychonomic Science*, *3*, 429–430.
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.
Paller, K. A., Bozic, V. S., Ranganath, C., Grabowecky, M., & Yamada, S. (1999). Brain waves following remembered faces index conscious recollection. *Cognitive Brain Research*, *7*, 519–531.
PsycheTruth (July 2009). How To Clean A Fish Tank or Aquarium, Remove Algae, Pet Care. Retrieved 30 September 2009 from http://www.youtube.com/watch?v=-yK7y0dKv00
Quinlan, C. K., Taylor, T. L., & Fawcett, J. M. (2010). Directed forgetting: Comparing pictures and words. *Canadian Journal of Experimental Psychology*, *64*, 41–46.
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*, 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*, 1064–1072.
Sahakyan, L., Waldum, E. R., Benjamin, A. S., & Bickett, S. (2009). Where is the forgetting with list method directed forgetting in recognition? *Memory & Cognition*, *37*, 464–476.
Taylor, T. L. (2005). Inhibition of return following instructions to remember and forget. *The Quarterly Journal of Experimental Psychology*, *58A*, 613–629.
Taylor, T. L., & Fawcett, J. M. (2012). Does an instruction to forget enhance memory for other presented items? *Consciousness & Cognition*, *21*, 1186–1197. doi:10.1016/j.concog.2012.05.002
Tekcan, 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*(6), 924–940.
Van Hooff, J. C., & Ford, R. M. (2011). Remember to forget: ERP evidence for inhibition in an item method directed forgetting paradigm. *Brain Research*, *1392*, 80–92.
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., Radvansky, G., & Hasher, L. (1996). Studies of directed forgetting in older adults. *Journal of Experimental Psychology: Learning, Memory, Cognition*, *22*, 143–156.
Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*, 3–21.