



# How prior testing impacts misinformation processing: A dual-task approach

Leamarie T. Gordon<sup>1</sup> · Vivek K. Bilollikar<sup>2</sup> · Taylor Hodhod<sup>3</sup> · Ayanna K. Thomas<sup>4</sup>

© The Psychonomic Society, Inc. 2019

## Abstract

Research suggests that testing prior to the presentation of misinformation influences how that misinformation is processed. The present study examined the relationship between testing, the demands of misinformation narrative processing, and memory for original and post-event information. Using response latencies to a secondary task, we tested whether prior testing influenced the available resources for secondary task processing. Additionally, we investigated whether changes in narrative processing were specific to critical details tested earlier. Participants engaged in an eyewitness memory paradigm in which half were tested prior to receiving the post-event narrative. Participants responded to the secondary task at specified time points during the narrative. All participants took a final memory test after listening to the post-event narrative. We found that testing interacted with the placement of the secondary task. When responding on the secondary task was closely linked to the presentation of previously tested critical details in the narrative, retrieval-enhanced suggestibility was reduced on tests of event memory (Experiment 1) and increased post-event information learning was revealed on tests of narrative memory (Experiment 2).

**Keywords** Misinformation effect · Retrieval enhanced suggestibility · Interim testing · Processing

## Introduction

When eyewitnesses are tested about a witnessed event, the act of memory retrieval or even attempted retrieval has been shown to influence attention to information presented after the test. That is, when presented with a post-event synopsis of the original event, research has demonstrated that participants spend more time reading details that are inconsistent (e.g., misleading details) with the original event (Gordon & Thomas, 2014; Gordon, Thomas, & Bulevich, 2015). These effects on reading time suggest changes in underlying attention processes. In the present study, we further examined how testing, or retrieval of event details, interacts with subsequent

attentional processes associated with processing of post-event information.

## Attention and retrieval enhanced suggestibility

Eyewitness memory has long been studied within the context of the misinformation paradigm (e.g., Loftus, Miller, & Burns, 1978). In a typical misinformation experiment, participants witness an original event, such as a video depicting a fictionalized crime, and after some retention interval are exposed to misleading post-event information about the witnessed event. The misinformation is often introduced in a post-event narrative or embedded within suggestive questions. Following the post-event information phase, memory for the witnessed event is measured. Decades of research employing this general procedure has revealed that exposure to misleading post-event information leads to errors of omissions, or a decreased likelihood of reporting accurate event details, and errors of commission, or an increased likelihood of erroneously reporting the suggested post-event details (Frenda, Nichols, & Loftus, 2011 for review).

Efforts have been made to increase the ecological validity of the misinformation paradigm. For example, Chan, Thomas, and Bulevich (2009) introduced a memory retrieval phase immediately following the witnessed event, which simulated

✉ Leamarie T. Gordon  
lt.gordon@assumption.edu

<sup>1</sup> Department of Psychology, Assumption College, Worcester, MA, USA

<sup>2</sup> Sidney Kimmel Medical College, Philadelphia, PA, USA

<sup>3</sup> Harvard Law School, Cambridge, MA, USA

<sup>4</sup> Department of Psychology, Tufts University, 490 Boston Ave, Medford, MA 02155, USA

the situation most eyewitnesses find themselves in when discussing their experience with first responders or other eyewitnesses at the scene of a crime. Interestingly, Chan et al. observed that when participants took this interim test between the witnessed event and the post-event information phase, the typical memory errors observed in the misinformation paradigm were even greater. That is, participants who took the interim test were less likely to recall accurate video information (e.g., errors of omission), and more likely to report suggested details from the post-event narrative (e.g., errors of commission) compared to participants who did not take the interim test. Subsequent studies have replicated these findings, and collectively dubbed them *retrieval-enhanced suggestibility* (RES; Chan et al., 2009; Chan & LaPaglia, 2013; Thomas, Gordon, Cernasov, & Bulevich, 2017; Thomas, Bulevich, & Chan, 2010).

Research in this area suggests that interim test questions in the misinformation paradigm may guide participants to differentially focus on test-relevant misinformation presented during the post-event narrative. For example, participants who took an interim memory test spent more time reading post-event narrative sentences containing misinformation compared to participants who did not take the interim test (Gordon & Thomas, 2014; Gordon et al., 2015). Gordon et al. (2015) further observed that when participants who took an interim test reported suggested misleading details on the final cued recall test, they had spent more time processing the misleading narrative sentences that introduced those details as compared to trials where they reported some other wrong answer on the final test. This difference was not present in participants who did not take an interim test. Finally, Gordon and Thomas (2017) found that when participants' ability to spend additional effort processing post-event misinformation following interim testing was disrupted, the RES effect was diminished. Thomas and colleagues have argued that attention shifts to narrative content after interim testing may in turn influence the ease with which misleading details are retrieved from memory on later tests (Gordon & Thomas, 2014; Gordon et al., 2015; Gordon & Thomas, 2017; Thomas et al., 2017). Indeed, studies designed to decrease the influence of temporary accessibility on memory retrieval have diminished RES (e.g., Thomas et al., 2010; Thomas et al., 2017). Interestingly, this body of research is in stark contrast to earlier findings that dividing attention during misinformation processing resulted in an increase in misinformation susceptibility in a standard misinformation paradigm (cf., Lane, 2006). The primary difference between the earlier and later work is the introduction of a test prior to misinformation presentation. In the present study, we examined how testing prior to narrative presentation interacted with attention manipulations to influence memory for the original event (Experiment 1) and memory for the post-event narrative (Experiment 2).

We argue that in the context of the RES paradigm, where a test prior to the misinformation is introduced, top-down attention guidance (e.g., McCormick, 1997; Stolz, 1996; Theeuwes, 1991; Yantis & Jonides, 1984) may be initiated when the narrative is processed. Information in the narrative may result in a search of memory for the previous test question as well as memory for one's response to the question. This shift in focus may guide future learning episodes and draw attention to post-event information that could be useful to answer the question, or corroborate or contradict the question response (cf. Chun & Turk-Browne, 2007). In fact, we have found in studies employing the RES paradigm that post-event narrative details were differentially processed (Gordon & Thomas, 2014; Gordon et al., 2015) and more fluently accessed (Thomas et al., 2010) on later retrieval attempts, resulting in RES errors of commission, or better memory for the narrative itself (Gordon & Thomas, 2017).

Our recent work has focused on how processing of the narrative changes as a result of prior testing. We have found that participants spend more time reading sentences that include details previously tested (e.g., Gordon & Thomas, 2014; Gordon et al., 2015). These longer reading times may reflect more attention to or more elaborative processing of these details than when an earlier test is absent. More recently, Gordon and Thomas (2017) disrupted participants' ability to covertly elaborate upon aurally introduced misleading post-event details by introducing a same-modality distractor task immediately following their presentation. The distractor task required participants to count a series of audio tones while simultaneously listening to the audio narrative. Under these constraints, Gordon and Thomas found that participants who took an interim test still reported misinformation presented in the narrative on a final memory test; however, the rates of these errors of commission were no worse than those observed in participants who did not take an interim test. The research reported by Gordon and colleagues consistently demonstrates that increased reports of misinformation, or narrative details, occur under conditions where participants engage in retrieval prior to narrative testing. We argue that prior retrieval may influence attention to relevant details in the narrative.

Based on this prior work, Thomas et al. (2017) predicted that non-retrieval-based attempts to manipulate attention to the post-event narrative may also elicit RES-like effects. That is, they predicted that when attention was drawn to critical details, participants would be less likely to report misinformation and more likely to report original event details. Indeed, they found that when participants encountered post-event misleading narrative details that were emphasized with a red, underlined font, RES errors of both omission and commission were similar to those observed in an interim test group who read the critical narrative details in a standard font. However, when a 48-h delay preceded the final memory test,

the manipulations revealed different impacts on memory. The interim testing group revealed better long-term memory accuracy for the original event compared to the emphasized narrative group, which is indicative of the relatively temporary influence that shifts in attention to misleading post-event information may have on eyewitness suggestibility. While Thomas et al. added an important piece to our understanding of how the distribution of processing resources may impact RES, they did not have any measure indicating that processing resources were comparably impacted by the interim testing and font-emphasis manipulations. Moreover, it is possible that grabbing attention exogenously by manipulating the physical attributes of a stimulus, such as the font, may engage a process distinct from that initiated by an interim test.

### The present study

The central objective of the present study was to better understand how interim testing influences processing of post-event information in an eyewitness memory paradigm. While Gordon and Thomas (2017) indirectly manipulated the availability of processing resources to observe its impact on RES, direct measurement of the potential shift in resource allocation following interim testing is presently limited to reading time studies using written post-event narratives (Gordon & Thomas, 2014; Gordon et al., 2015). Thus, we aimed to provide converging evidence by measuring attention during processing of an aural, experimenter-paced narrative. To do this we capitalized on the boundaries of the limited-capacity cognitive system.

According to the limited-capacity model of attention (e.g., Kahneman, 1973; Lynch & Srull, 1982), the total attentional capacity needed to perform simultaneous activities can be divided into two parts: capacity devoted to a primary task and spare capacity to complete secondary tasks. When the demands of the primary task increase, distribution of resources across these categories becomes less equitable. Moreover, whether the secondary task will impact performance on the primary task is dependent on the structure of the secondary task (Wickens, 1980). Wickens's (2008) multidimensional resource allocation model defines "resources" according to three dimensions: stage of processing (perception/cognition or response), code of processing (spatial or verbal), and modality (visual or auditory). The model predicts that the more dimensions shared by concurrent tasks, the worse the performance. For example, simultaneous performance on a visual-verbal task and an auditory-verbal task (which share the code of processing dimension) should be worse than performance on a visuospatial task and an auditory-verbal task (which differ along said dimension). In other words, some secondary tasks will interfere with primary tasks more than others. In Gordon and Thomas (2017), participants' ability to spend additional time processing just-encoded narrative details (the *primary*

task) was successfully disrupted by a secondary task sharing multiple processing dimensions. However, even very modest amounts of overlap across processing pathways can constrain how many cognitive processes can be executed simultaneously (e.g., Shenav et al., 2017), making a carefully constructed divided attention task a practical way to measure resource allocation during post-event narrative processing in the RES paradigm. Specifically, when participants are engaged with a primary task, such as encoding aurally presented narrative details, it follows that performance on an unrelated secondary task, such as making a time-sensitive response to a visual probe presented on a computer screen, will depend on how much "spare capacity" is available to complete the probe task. When manipulations that have been shown to increase the demands of the primary narrative task are employed, such as preceding testing, the spare capacity available for the secondary task should reduce even further.

Thus, in the present study we measured reaction times to an unrelated secondary task presented during post-event narrative encoding in order to characterize the availability of spare resources. Immediately following the introduction of critical narrative details, where processing resources are presumed to be occupied with elaborative encoding (Gordon & Thomas, 2017), participants were presented with a different modality visual probe on a computer screen and pressed the appropriate key depending on the probe's location. The critical probe was presented 1,500 ms after participants heard a critical detail. This time frame was chosen because it captured the average reading time window of sentences that included critical details reported in prior research (see Gordon & Thomas, 2014). Response times to the probe were recorded. We hypothesized that participants would respond more slowly to probes presented after misleading narrative details compared to neutral or consistent narrative details because they would be spending more time processing misleading details.

We also used the probe task to better understand the parameters of test-induced shifts in attention in the RES paradigm. Gordon et al. (2015) previously proposed that taking an interim test would change participants' overall strategy during narrative reading (see also Britton, Piha, Davis, & Wehausen, 1978; Reynolds, Standiford, & Anderson, 1979), which would result in group differences in processing time across all item types. Gordon et al. (2015) demonstrated that while the standard group spent a statistically equivalent amount of time processing narrative sentences of each type (e.g., misleading, neutral, and consistent), the interim test group spent significantly more time processing consistent and misleading sentences compared to neutral sentences. That is, in the context of full attention, interim testing changed how participants distributed attention, at least across sentences containing critical test-relevant information. It is presently unknown whether this change in attention extends globally across the entire narrative, including non-critical information, or is restricted to

critical items only. Thus, in the present study we included a second probe group who experienced the same dual-task manipulation as the first, except the probes were presented 6 s before or after critical details, effectively disassociating the probe from critical detail processing. We hoped this manipulation would give us some insight into the temporal boundaries of test-induced attention shifts. In Experiment 1, the final test measured memory for the video. In Experiment 2, the final test measured memory for the post-event narrative.

## Experiment 1

### Method

#### Design

The study employed a 2 (Testing Group: Interim Test, Standard)  $\times$  3 (Probe Placement: Critical, Non-Critical, No Probe)  $\times$  3 (Item Type: Consistent, Neutral, Misleading) mixed design. Testing Group and Probe Placement were manipulated between subjects while Item Type was manipulated within subjects.

#### Participants

A sample size estimation was calculated using G\*Power Version 3 software. Using moderate parameters (power = 0.8, effect size = 0.25), the analysis estimated a minimum sample size of 144. A total of 240 undergraduates from Tufts University participated and were compensated either with course credit or were paid US\$15. Participants were drawn from a pool of participants that consisted of 65% of individuals who identified as women, 33% who identified as men, and 2% who identified as non-binary or chose to not provide this information. All participants were between the ages of 18 and 24 years.

#### Materials and procedure

Participants were tested in small groups no larger than four. Participants were seated at personal work stations, and other than group instructions, experienced the study individually. All participants first watched a 22-min video clip of the black and white silent film “Rififi” (Bezard, Bérard, Cabaud, & Dassin, 1955). The clip portrayed a group of four men committing a burglary in the middle of the night. No participant had viewed the video before.

Following the video, participants in the interim test group took an immediate cued recall test on details from the video. Twenty-four questions, taken from Gordon and Thomas (2017), were used as interim test stimuli. Each question targeted critical details manipulated in the upcoming narrative

(e.g., *What did the burglar remove from the drawer of valuables?*), and were presented in the same order as information was encountered during the video. Participants viewed each question individually on a computer screen, and had 15 s to type in their response. After 15 s, the program would automatically advance to the next question. Participants were encouraged to answer each question, but were able to withhold responses. Instead of taking the interim test, the standard group completed a Sudoku puzzle for an equivalent amount of time (6 min). Participants in both groups then completed filler tasks that included a brief demographic questionnaire, and a synonym and antonym vocabulary test. The responses on the filler tasks were not recorded for use in any analysis. Next, all participants listened to an audio recording of a female narrator describing the events from the video. The narrative was presented at typical conversational pace. Critically, the narrative contained 24 sentences that introduced consistent, neutral, and misleading information about the video (eight details each), in addition to filler information that was not tested in either the interim or final test phases (also from Gordon & Thomas, 2017). Consistent sentences contained details that were accurate regarding the encoding event (e.g., *Before exiting, he tosses an envelope into the safe.*) Neutral information was a detail presented in the video, but not manipulated in the narrative (e.g., *Before exiting, he tosses something into the safe.*) Misleading information always consisted of a detail in the video that had been changed in the narrative (e.g., *Before exiting, he tosses his gloves into the safe.*) Sentences serving as misleading, neutral, and consistent were counterbalanced across participants.

During the narrative presentation, participants in each testing group had one of three different experiences. Two of these experiences required participants to view a computer screen while listening to the narrative. Periodically throughout the narrative, an on-screen cue (a non-verbal symbol:  $\leftarrow$ ) appeared alerting these participants to make a keyboard response. Reaction times for these responses were recorded. If a cue appeared on the left side of the screen, participants were instructed to press a key on the right-hand side of the keyboard. If a cue appeared on the right side of the screen, participants were instructed to press a key on the left-hand side of the keyboard.

The temporal placement of these simple perceptual judgments, or probe trials, varied across the experiences. For participants in the Critical Probe group, probes appeared temporally close to critical narrative details. Specifically, the probes appeared 1,500 ms after the critical detail in each sentence was presented. Critical details were presented at any location within a sentence. Gordon and Thomas (2017) demonstrated that disrupting processing 1,500 ms following initial encoding of critical details in this paradigm minimized retrieval-enhanced suggestibility, which suggests that this is a critical period for elaborative processing of those details. For participants in the



Non-Critical Probe group, probe presentation was not temporally tied to critical narrative details. Instead, the probe was randomly presented either 6 s before or 6 s after critical narrative details, and always in the context of a filler narrative sentence containing no test-relevant information. In the third narrative experience, the No Probe group listened to the narrative without any visual probe presentation and thus did not make keyboard responses. Importantly, the only difference between the two probe placement groups was the distance from a critical detail.

Following the narrative phase, all participants took a final cued recall memory test that was identical to the interim test completed by the interim test group. Participants were instructed to respond only with information they remembered from the original encoding event (video). Participants were then thanked and debriefed.

## Results

### Reaction time to probe task

The reaction-time analyses were limited to trials on which the response to the probe was correct. On average, participants responded correctly to the probe when presented after consistent details 91% of the time, after neutral details 90% of the time, and after misleading details 89% of the time. In addition, any reaction time longer than 4,000 ms was removed from analysis, as this indicated non-compliance with task instructions (this occurrence was rare). The No Probe group was not included in these analyses as they did not perform the probe task.

A 2 (Testing Group: interim, standard)  $\times$  2 (Probe Placement: critical, non-critical)  $\times$  3 (Item Type: consistent, neutral, misleading) mixed ANOVA examined mean reaction times on the probe task. The main effect of Item Type was significant,  $F(2, 310) = 4.28, p < .05, \eta_p^2 = .03$ . Overall, participants took longer to respond to probes presented on misleading trials ( $M = 753\text{ms}$ ) compared to probes presented on neutral trials ( $M = 723\text{ms}$ ),  $t(158) = 2.26, p < .05, d = .18$ . Participants also took longer to respond to probes on misleading compared to consistent questions ( $M = 722\text{ms}$ ),  $t(158) = 2.85, p < .01, d = .23$ . The difference between response times on neutral and consistent trials was not significant,  $p > .05$ .

In addition, the Testing Group  $\times$  Probe Placement interaction was significant,  $F(1, 155) = 4.60, p < .05, \eta_p^2 = .03$ . Independent groups  $t$  tests conducted separately within each probe group followed up with this interaction. These  $t$  tests were corrected for alpha inflation using the Bonferroni correction. In the critical probe group, participants who took the interim test ( $M = 694\text{ms}$ ) responded more quickly to probes than participants in the standard group ( $M = 768\text{ms}$ ). Although the omnibus test revealed a significant interaction,

this specific comparison was marginally significant,  $t(78) = 1.83, p = .07$ . In the non-critical probe group, the difference in probe reaction times did not differ between the interim test ( $M = 751\text{ms}$ ) and standard ( $M = 716\text{ms}$ ) groups,  $p = .35$ . No other effects were significant. Table 1 presents mean reaction times for each cell.

### Memory performance

All follow-up comparisons used a Bonferroni correction unless otherwise stated. Accurate recall was calculated by dividing the total number of trials in which participants produced the correct video detail out of the total number of trials for that given item type. During the interim recall test, .54 of participants' responses were accurate and .06 produced misinformation spontaneously. Although we did not expect probe placement to impact memory performance, we included that variable in all final test performance analyses.

**Accurate video recall on final test** Table 2 presents the accurate recall probabilities on the final test. A 2 (Testing Group: interim, standard)  $\times$  3 (Probe Placement: critical, non-critical, no probe)  $\times$  3 (Item Type: consistent, neutral, misleading) mixed analysis of variance (ANOVA) examined final accurate recall. It revealed a main effect of Item Type,  $F(2, 468) = 332.26, p < .001, \eta_p^2 = .59$ . Participants were most accurate on consistent questions ( $M = .81$ ), followed by neutral items ( $M = .59$ ), and finally misleading question trials ( $M = .41$ ). Follow-up tests revealed that participants recalled more accurate details on consistent trials compared to neutral trials,  $t(239) = 15.07, d = .97$ . In addition, participants were more accurate on neutral trials as compared to misleading trials,  $t(239) = 11.29, d = .73$ .

The ANOVA on accurate recall also revealed an Item Type  $\times$  Testing Group interaction,  $F(2, 468) = 10.23, p < .001, \eta_p^2 = .04$ . On consistent trials, the interim test group ( $M = .84$ ) was more accurate than the standard group ( $M = .79$ ),  $t(238) = 2.13, p < .05, d = .27$ . However, on misleading trials the interim test group ( $M = .37$ ) recalled fewer accurate video details than the standard group ( $M = .46$ ), establishing a

**Table 1** Average reaction times to probe task in milliseconds in Experiment 1 (standard error in parentheses)

	Consistent	Neutral	Misleading
Critical Probe			
Standard	739 (30)	783 (31)	782 (30)
Interim Test	674 (30)	687 (31)	719 (30)
Non-Critical Probe			
Standard	712 (30)	711 (31)	726 (30)
Interim Test	764 (31)	728 (31)	787 (31)

**Table 2** Average proportion of video details recalled on the final test in Experiment 1 (standard error in parentheses)

	Consistent	Neutral	Misleading
Critical Probe			
Standard	.81 (.03)	.58 (.03)	.44 (.04)
Interim Test	.85 (.03)	.61 (.03)	.39 (.04)
Non-Critical Probe			
Standard	.80 (.03)	.64 (.03)	.48 (.04)
Interim Test	.85 (.03)	.61 (.03)	.40 (.04)
No Probe			
Standard	.76 (.03)	.61 (.03)	.45 (.04)
Interim Test	.84 (.03)	.56 (.03)	.34 (.04)

retrieval enhanced suggestibility effect for recall accuracy,  $t(239) = 3.15$ ,  $p < .01$ ,  $d = .40$ . The difference between testing groups on neutral trials was not significant,  $p = .31$ . No other main or interaction effects were significant,  $ps > .05$ .

**Misinformation production on final test** Table 3 reports probabilities of errors of commission, or misinformation production, on the final test. Misinformation production was calculated by dividing the number of trials in which participants produced the misleading detail out of the total number of trials for each item type.

A 2 (Testing Group: interim, standard)  $\times$  3 (Probe Placement: critical, non-critical, no probe)  $\times$  3 (Item Type: consistent, neutral, misleading) mixed ANOVA analyzed misinformation production. It revealed a main effect of item type,  $F(2, 468) = 491.44$ ,  $p < .001$ ,  $\eta_p^2 = .68$ . Overall, participants reported misinformation on .44 of misleading question trials. Spontaneous reports of misinformation on consistent (.05) and neutral (.09) question trials were rare. Establishing a standard misinformation effect, the difference in misinformation reports between misleading and neutral trials was significant,  $t(239) = 20.48$ ,  $d = 1.32$ .

**Table 3** Average proportion of misleading narrative details reported on the final test in Experiment 1 (standard error in parentheses)

	Consistent	Neutral	Misleading
Critical Probe			
Standard	.05 (.01)	.12 (.02)	.40 (.04)
Interim Test	.05 (.01)	.11 (.02)	.44 (.04)
Non-Critical Probe			
Standard	.08 (.01)	.10 (.02)	.35 (.04)
Interim Test	.04 (.01)	.10 (.02)	.55 (.04)
No Probe			
Standard	.06 (.01)	.13 (.02)	.36 (.04)
Interim Test	.04 (.01)	.07 (.02)	.53 (.04)

A main effect of Testing Group was also significant,  $F(1, 234) = 7.52$ ,  $p < .01$ ,  $\eta_p^2 = .03$ . Consistent with the RES literature, interim test participants reported more misinformation ( $M = .21$ ) than standard participants ( $M = .18$ ). We also found an Item Type  $\times$  Testing Group interaction,  $F(2, 468) = 23.75$ ,  $p < .001$ ,  $\eta_p^2 = .09$ . Interim test participants ( $M = .04$ ) were less likely than the standard group ( $M = .06$ ) to spontaneously guess misleading details on consistent trials,  $t(238) = 2.11$ ,  $p < .05$ ,  $d = .28$ . More relevant is the comparison on misleading trials. Here, the interim test group was more likely to respond with misleading details presented in the narrative ( $M = .51$ ) compared to the standard group ( $M = .37$ ),  $t(239) = 4.70$ ,  $p < .001$ ,  $d = .61$ . The difference between testing groups on neutral trials was not significant,  $p = .10$ .

Finally, a Testing Group  $\times$  Probe Placement  $\times$  Item Type interaction was significant,  $F(4, 468) = 3.43$ ,  $p < .01$ ,  $\eta_p^2 = .03$ . To unravel this interaction, a 2 (Testing Group: interim, standard)  $\times$  3 (Item Type: consistent, neutral, misleading) mixed-factor ANOVA was conducted separately in each probe placement group. In the no-probe group RES was revealed by the significant Testing Group  $\times$  Item Type interaction,  $F(2, 156) = 13.68$ ,  $p < .01$ ,  $\eta_p^2 = .15$ . Importantly, participants in the interim test group ( $M = .53$ ) reported more misleading details than the standard group ( $M = .36$ ),  $t(78) = 3.38$ ,  $p < .01$ ,  $d = .77$ . In the non-critical probe group RES was also revealed by a significant Testing Group  $\times$  Item Type interaction,  $F(2, 156) = 17.96$ ,  $p < .01$ ,  $\eta_p^2 = .19$ . Importantly, participants in the interim test group ( $M = .55$ ) reported more misleading details than the standard group ( $M = .35$ ),  $t(78) = 4.28$ ,  $p < .01$ ,  $d = .95$ . Finally, RES was *not* observed in the critical probe group as the Testing Group  $\times$  Item Type interaction was not significant,  $p = .48$ . That is, although a typical misinformation effect occurred in this group, interim test participants reported on average the same number of misleading details on the final memory test compared to the group who did not take an interim test (see Table 3).

## Discussion

The primary goal of the first experiment was to directly measure the availability of spare attentional resources during narrative processing in the RES paradigm. Toward this end, participants responded to a visual probe presented at specific time points during the aural narrative and reaction times served as an indication of available resources during processing of the narrative. We found that participants had the slowest response times on misleading narrative trials, suggesting that participants' attention was drawn to narrative details that contradict an originally witnessed event. We also found that probe placement mattered. Participants in the critical probe group responded faster to the probe after taking an interim test;

however, this effect was relatively small, and we planned to replicate it in Experiment 2. Regarding the final memory test, typical RES errors of omission and commission were observed, with the interim test participants recalling fewer correct details, and producing more suggested details on misleading trials compared to standard participants. In addition, probe placement interacted with testing group. That is, participants who took an interim test and then engaged in the critical probe manipulation were similarly likely to produce misleading details as the standard group.

The critical probe group yielded faster response times to the probe on misleading trials after initial testing, no change in memory accuracy on misleading trials, and a reduction in misinformation production on misleading trials when compared to the other conditions. This pattern of results was somewhat unexpected. Faster response times to the probe coupled with a reduction in misinformation production suggests that placing probes in the same temporal context as critical narrative information may have disrupted processing of those critical details. We developed the secondary task to have little impact on primary-task processing and completion; however, the possibility that primary-task processing was affected remains. In addition, given that the reaction-time effect was rather small, and the findings contrast our initial predictions, we developed a follow-up experiment.

## Experiment 2

The primary goals of Experiment 2 were to replicate the reaction-time results found in Experiment 1 and to further test whether processing of critical details was disrupted by the secondary critical probe task. Although the bulk of misinformation experiments focus on memory for the original event, Experiment 2 assessed memory for the narrative. Perhaps less interesting to eyewitness memory researchers, memory for post-test information has theoretical significance as it informs our understanding of how testing influences processing of information after the test. In fact, interim testing has been shown to result in forward effects of testing, in that memory for information presented after the interim test is enhanced (e.g., Gordon & Thomas, 2017). In Experiment 2, participants were instructed to respond only with information learned in the narrative on the final memory test. If the secondary probe task disrupts processing of critical details, fewer details should be reported on a test specifically targeting memory for those details. Further, this manipulation allowed us to test whether interim testing and the critical probe task interacted to influence memory for the post-test material.

## Method

### Design

Experiment 2 employed a 2 (Testing Group: interim test, standard)  $\times$  2 (Probe Placement: critical, non-critical)  $\times$  3 (Item Type: consistent, neutral, misleading) mixed design. As we were interested in the influence of probe placement on narrative learning, we did not include a no-probe group. As in Experiment 1, testing group and probe placement were manipulated between participants while item type was manipulated within participants.

### Participants

A total of 200 undergraduates from Tufts University participated and were compensated either with course credit or were paid US\$15. The sample size for this modified design was determined using the same parameters as Experiment 1. An equal number of participants were randomly assigned to four between-participants groups. Participants were drawn from the same participant pool as in Experiment 1.

### Materials and procedure

The materials and procedure used in Experiment 2 were the same as in Experiment 1 with two exceptions. First, we did not include a no-probe group, as this group was not necessary to examine any of the present hypotheses. Second, on the final memory test participants were instructed to report what they remembered learning in the narrative only.

## Results

### Reaction time to probe task

As in Experiment 2, the reaction-time analyses were limited to trials on which the response to the probe was correct. On average, participants responded correctly to 87% of the probes presented in association with critical details. There was no difference in accuracy of response rate as a function of trial type. In addition, any reaction time longer than 4,000 ms was removed from analysis, as this indicated non-compliance with task instructions (this occurrence was rare).

A 2 (Testing Group: interim, standard)  $\times$  2 (Probe Placement: critical, non-critical)  $\times$  3 (Item Type: consistent, neutral, misleading) mixed ANOVA examined mean reaction times on the probe task. The main effect of Item Type was significant,  $F(2, 392) = 5.87, p < .003, \eta_p^2 = .03$ . Overall, participants took longer to respond to probes presented on misleading trials ( $M = 736$  ms) compared to probes presented on neutral trials ( $M = 717$  ms),  $t(199) = 2.04, p < .05, d = .09$

(although this difference was not significant after a Bonferroni correction). Participants also took longer to respond to probes on misleading compared to consistent trials ( $M = 708$  ms),  $t(199) = 3.57$ ,  $p < .001$ ,  $d = .14$ . The difference between response times on neutral and consistent trials was not significant,  $p > .05$ .

In addition, the Testing Group  $\times$  Probe Placement interaction was significant,  $F(1, 196) = 4.11$ ,  $p < .05$ ,  $\eta_p^2 = .02$ . Independent groups  $t$  tests conducted separately within each probe group followed up with this interaction. These  $t$  tests were corrected for alpha inflation using the Bonferroni correction. In the critical probe group, participants who took the interim test ( $M = 653$  ms) responded more quickly to probes than participants in the standard group ( $M = 731$  ms),  $t(98) = 2.87$ ,  $p < .005$ ,  $d = .60$ . In the non-critical probe group, the difference in probe reaction times did not differ between the interim test ( $M = 763$  ms) and standard ( $M = 736$  ms) groups,  $p = .56$ . No other effects were significant. Table 4 presents mean reaction times for each cell.

### Memory performance

During the interim recall test of the original event, .57 of participants' responses were accurate and .05 produced misinformation spontaneously.

**Narrative recall** On the final test, participants were instructed to report information they learned in the *narrative only*. For consistent trials, narrative details corroborated those presented in the video. For misleading trials, narrative details were new, and contradicted those presented in the video. Neutral trials did not provide specific details in the narrative so were excluded from analysis. Bonferroni corrections were used where appropriate.

A 2 (Testing Group: interim, standard)  $\times$  2 (Probe Placement: critical, non-critical)  $\times$  2 (Item Type: consistent, misleading) mixed ANOVA was conducted on accurate recall of narrative details. A main effect of Item Type was significant,  $F(1, 196) = 72.57$ ,  $p < .001$ ,  $\eta_p^2 = .27$ . As Table 5 demonstrates, participants recalled more narrative details when

**Table 5** Average proportion of narrative details recalled the final test in Experiment 2 (standard error in parentheses)

	Consistent	Misleading
Critical Probe		
Standard	.66 (.03)	.43 (.04)
Interim Test	.82 (.03)	.75 (.03)
Non-Critical Probe		
Standard	.67 (.04)	.42 (.04)
Interim Test	.79 (.04)	.66 (.04)

they were presented as consistent with the video ( $M = .73$ ) compared to when they contradicted the video ( $M = .57$ ). A main effect of Testing Group was also found,  $F(1, 196) = 82.13$ ,  $p < .001$ ,  $\eta_p^2 = .29$ . Participants in the interim test group ( $M = .75$ ) recalled more narrative details than participants in the standard group ( $M = .55$ ). A main effect of Probe Placement was also found,  $F(1, 196) = 4.41$ ,  $p < .05$ ,  $\eta_p^2 = .02$ . However, this main effect should be considered in the context of the interaction between Testing Group and Probe Placement,  $F(1, 196) = 5.36$ ,  $p < .05$ ,  $\eta_p^2 = .12$ . Follow up  $t$  tests using a Bonferroni correction found that whereas probe placement had no effect on memory in the standard group, participants in the critical probe group demonstrated better memory for the narrative if they had taken an interim test,  $t(98) = 2.94$ ,  $d = 0.56$ . Finally, the Testing Group  $\times$  Item Type interaction was significant,  $F(1, 196) = 12.90$ ,  $p < .001$ ,  $\eta_p^2 = .06$ . As Table 5 demonstrates, participants in the interim test group reported more critical narrative details than those in the standard group; however, the difference was greater on misleading as compared to consistent trials. These data are consistent with those reported by Gordon and Thomas (2017), and suggest that interim testing has forward effects that impact learning of narrative details.

### Discussion

Results from Experiment 2 replicate and extend the findings reported in Experiment 1. Testing paired with critical probe placement affected attention and processing of critical event details. These changes in processing had consequences for reaction times to the probe, as well as memory for narrative details. These results suggest that interim testing does influence post-test narrative learning, as demonstrated by performance on the final memory test. These results also suggest the critical probe task interacted with interim testing. We argue in the *General discussion* that this interaction resulted in enhanced memory for items in the narrative, and memory for where the items were presented (e.g., video or narrative).

**Table 4** Average reaction times to probe task in milliseconds in Experiment 2 (standard error in parentheses)

	Consistent	Neutral	Misleading
Critical Probe			
Standard	706 (13)	727 (15)	761 (23)
Interim Test	651 (25)	636 (22)	674 (26)
Non-Critical Probe			
Standard	722 (32)	754 (34)	731 (32)
Interim Test	755 (31)	755 (35)	780 (34)



## General discussion

The main goal of the present study was to better understand how interim testing may influence processing of post-event information in an eyewitness memory paradigm. Some prior studies have suggested that interim test questions can guide participants to differentially process narrative content relevant to the questions (Gordon & Thomas, 2014; Gordon et al., 2015). Assuming that additional processing of narrative content requires additional cognitive resources, we further explored the nature of these processing changes by measuring the time it took participants to respond to an unrelated visual probe task at specified points during the post-event narrative presentation. This design gave us a novel way to examine the availability of spare resources after interim testing.

Consistent with prior RES studies (Butler & Loftus, 2018; Chan et al., 2009; Gordon et al., 2015; LaPaglia, Wilford, Rivard, Chan, & Fisher, 2014), participants who took an interim test were less likely to report accurate video details, and when they were wrong, more likely to report suggested narrative details on the final memory test compared to participants who did not take an interim test. However, placement of the probe in the context of the narrative affected these memory findings. In Experiment 1, when participants' memory for the original event was tested on the final memory test, we found an interaction between probe placement and testing group. Interim test participants demonstrated typical RES effects when processing the narrative in isolation (no probe) and in the context of the non-critical probe condition. That is, these groups of participants were more likely to produce misleading details on the final test than participants who did not take the interim test. However, when a probe was temporally closer to the presentation of a critical detail, participants in the interim test group were no more likely to produce misleading details than participants in the standard group. While they demonstrated a standard misinformation effect, the absence of an RES production effect in this group suggests that the probe task may have disrupted additional processing after testing, similar to the divided attention task used in Gordon and Thomas (2017).

In Gordon and Thomas, the counting task used to divide participants' attention following encoding of critical narrative details led to two important findings. First, the secondary task reduced retrieval-enhanced errors of commission when video memory was tested. Second, it impaired memory for misleading details across both testing groups when narrative memory was tested. In the present study, we were able to examine how the probe response task impacted narrative memory by instructing participants in Experiment 2 to respond only with information learned in the narrative. We found here that participants in the interim test group demonstrated better memory for narrative details than participants in the standard group. Not only that, but in contrast to Gordon and Thomas (2017)

we also found that when the probe task followed encoding of critical narrative details, there was no indication of memory impairment for those details compared to the non-critical probe group. In fact, narrative memory in the critical probe group improved following interim testing. This suggests a difference in the processing demands required between the secondary task used in Gordon and Thomas and the probe detection task used in the present study. It is important to note that the present Experiment 2 did not include a no-probe group, which limits direct comparisons to Gordon and Thomas. To summarize the memory performance results from the two present experiments, it appears that interim testing potentiates learning of the narrative. In addition, when coupled with a task that draws attention to previously tested details, learning of those details improves. These improvements were demonstrated in a reduction in misinformation production errors (Experiment 1) and improved ability to recollect narrative details (Experiment 2) when compared to participants who did not take an interim test, and to participants who engaged in a secondary task less temporally associated with critical detail processing.

The primary motivation for using the probe task was to see if response times could provide insight into how participants may be processing narrative details. In both experiments, participants were slowest to respond to probes tied to misleading narrative details. This finding suggests that independent of preceding testing, misinformation may draw more attentional resources, resulting in a cost to responding to a secondary probe (cf. Gordon et al., 2015). Based on previous findings that interim testing results in slower processing of test-relevant narrative details (Gordon & Thomas, 2014), we had predicted that critical probe placement would result in slower response times, and this slow down would be more pronounced following testing. Probe placement consistently interacted with the presence of interim testing across experiments. However, the interaction indicated a pattern opposite to our prediction.

In both experiments, we found that when the visual probe was presented temporally close to the critical narrative detail, participants in the interim test group responded *more quickly* than participants in the standard group. Taken together, these results suggest that: (1) testing may facilitate processing of critical narrative details; (2) critical probes may serve to tag those details, affording interim test participants an additional cue that was useful when differentiating between details presented in the video from those presented in the narrative. This differentiation resulted in a reduction in misinformation production in Experiment 1, effectively eliminating RES. This differentiation also resulted in better memory performance on the final test of memory in Experiment 2. Although we did not set up these experiments to test alternative hypotheses, our findings suggest that a mechanism other than one that is capacity based may explain the interaction between probe placement and testing.

One *post hoc* explanation that requires additional confirmatory experimentation is that the visual probe may have been processed as an associated element of the critical detail. After interim testing, participants may temporarily spend extra effort elaborating upon narrative details or devote more attention to those details (cf. Gordon & Thomas, 2017). This change in processing may influence how participants process associated stimuli, such as a visual probe presented temporally close to a critical detail. A growing body of work has found that under some conditions attending to a primary encoding task can facilitate performance on a second, unrelated target detection task, a phenomenon known as the *attentional boost effect* (ABE; see Swallow & Jiang, 2013, for review). This effect has been consistently demonstrated with both visual and verbal encoding materials (Mulligan, Spataro, & Picklesimer, 2014; Spataro, Mulligan, & Rossi-Arnaud, 2011; Swallow & Jiang, 2010). ABE contrasts with the majority of the dual-task performance and attention literature by showing that increasing attention to one task can trigger an attentional process that supplements, rather than impairs, performance on a second task. According to this alternative explanation, the temporary increase in attention given to test-relevant details may have facilitated processing of the visual probe, resulting in quicker response times. Facilitated probe detection did not occur when probes were tied to non-critical narrative information in either experiment. This suggests that the probe detection “boost” that may occur during processing of critical narrative details is relatively brief and does not spill over to noncritical information. We caution that this explanation is preliminary at best; however, these findings certainly warrant further exploration. The present experiments were not designed to test hypotheses associated with ABE.

## Conclusion

The present experiments were designed to examine how interim testing influenced post-test processing of narrative details. Consistent with prior work in this area, the present study demonstrated that interim testing impacts how post-event narrative details are processed, and this had downstream consequences for memory performance. The present results have important implications for understanding the nature of eyewitness memory accuracy and susceptibility to post-event information. Counter to what some may intuitively suspect, repeated questioning can introduce errors. Repeated questioning may result in potentiation of learning of new details via attentional shifts towards that information. However, our results also suggest that potentiation of new learning does not overwrite old learning. Cues that may become associated with potentiated learning may help to distinguish those details from original event details.

**Open Practices Statement** None of the data or materials for the experiments reported here are currently available electronically, and none of the experiments were preregistered.

## References

- Bezard, R., Bérard, H., Cabaud, P., & Dassin, J. (1955). Riffifi [Motion Picture]. France, Pathé Consortium Cinéma.
- Britton, B. K., Piha, A., Davis, J., & Wehausen, E. (1978). Reading and cognitive capacity usage: Adjunct question effects. *Memory & Cognition*, 6(3), 266-273.
- Butler, B. J., & Loftus, E. F. (2018). Discrepancy detection in the retrieval-enhanced suggestibility paradigm. *Memory*, 26(4), 483-492.
- Chan, J. C. K., & LaPaglia, J. A. (2013). Impairing existing declarative memory in humans by disrupting reconsolidation. *PNAS: Proceedings of the National Academy of Sciences of the United States of America*, 110(23), 9309-9313.
- Chan, J. C. K., Thomas, A. K., & Bulevich, J. B. (2009). Recalling a witnessed event increases eyewitness suggestibility: The reversed testing effect. *Psychological Science*, 20, 66-73.
- Chun, M. M., & Turk-Browne, N. B. (2007). Interactions between attention and memory. *Current Opinion in Neurobiology*, 17, 177-184.
- Frenda, S. J., Nichols, R. M., & Loftus, E. F. (2011). Current issues and advances in misinformation research. *Current Directions in Psychological Science*, 20(1), 20-23.
- Gordon, L. T., & Thomas, A. K. (2014). Testing potentiates new learning in the misinformation paradigm. *Memory & Cognition*, 42(2), 186-197.
- Gordon, L. T., & Thomas, A. K. (2017). The forward effects of testing on eyewitness memory: The tension between suggestibility and learning. *Journal of Memory and Language*, 95, 190-199.
- Gordon, L. T., Thomas, A. K., & Bulevich, J. B. (2015). Looking for answers in all the wrong places: How testing facilitates learning of misinformation. *Journal of Memory and Language*, 83, 140-151.
- Kahneman, D. (1973). Attention and effort. Englewood Cliffs, NJ: Prentice-Hall, Inc.
- Lane, S. M. (2006). Dividing attention during a witnessed event increases eyewitness suggestibility. *Applied Cognitive Psychology*, 20(2), 199-212. <http://dx.doi.org.ezproxy.library.tufts.edu/10.1002/acp.1177>
- LaPaglia, J. A., Wilford, M. M., Rivard, J. R., Chan, J. C., & Fisher, R. P. (2014). Misleading suggestions can alter later memory reports even following a cognitive interview. *Applied Cognitive Psychology*, 28(1), 1-9.
- Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory*, 4(1), 19.
- Lynch, J. G., & Srull, T. K. (1982). Memory and attentional factors in consumer choice: Concepts and research methods. *Journal of Consumer Research*, 9, 18-37.
- McCormick, P. A. (1997). Orienting attention without awareness. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 168-180.
- Mulligan, N. W., Spataro, P., & Picklesimer, M. (2014). The attentional boost effect with verbal materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 40, 1049-1063.
- Reynolds, R. E., Standiford, S. N., & Anderson, R. C. (1979). Distribution of reading time when questions are asked about a restricted category of text information. *Journal of Educational Psychology*, 71(2), 183.

- Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a rational and mechanistic account of mental effort. *Annual Review of Neuroscience*, *40*, 99–124.
- Spataro, P., Mulligan, N. W., & Rossi-Arnaud, C. (2011). Attention and implicit memory. The role of the activation of multiple representations. *Experimental Psychology*, *58*, 110–116.
- Stolz, J. A. (1996). Exogenous orienting does not reflect an encapsulated set of processes. *Journal of Experimental Psychology: Human Perception and Performance*, *22*(1), 187.
- Swallow, K.M., & Jiang, Y.V. (2010). Attentional boost effect: Transient increases in attention to one task enhance performance in a second task. *Cognition*, *115*, 115–132.
- Swallow, K. M., & Jiang, Y. V. (2013). Attentional load and attentional boost: A review of data and theory. *Frontiers in Psychology*, *4*, 274.
- Theeuwes, J. (1991). Exogenous and endogenous control of attention: The effect of visual onsets and offsets. *Perception & Psychophysics*, *49*, 83–90.
- Thomas, A. K., Bulevich, J. B., & Chan, J. C. (2010). Testing promotes eyewitness accuracy with a warning: Implications for retrieval enhanced suggestibility. *Journal of Memory and Language*, *63*(2), 149–157.
- Thomas, A. K., Gordon, L. T., Cernasov, P. M., & Bulevich, J. B. (2017). The effect of testing can increase or decrease misinformation susceptibility depending on the retention interval. *Cognitive Research: Principles and Implications*, *2*(1), 45.
- Wickens, C. D. (1980). The structure of attentional resources. In R. Nickerson (Ed.), *Attention and performance VIII* (pp. 239–257). Hillsdale, NJ: Erlbaum.
- Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors*, *50*, 449–455.
- Yantis, S., & Jonides, J. (1984). Abrupt onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 601–620.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.