Spring 5-1-2017

The New Theory of Disuse Predicts Retrieval Enhanced Suggestibility (RES)

Victoria Bartek
victoria.bartek@student.shu.edu

The New Theory of Disuse Predicts Retrieval Enhanced Suggestibility (RES)

By Victoria Bartek

A Thesis Submitted In Partial Fulfillment of the Requirements for the Master of Science in Experimental Psychology with a Concentration in Cognitive Neuroscience

The Department of Psychology
Seton Hall University
May, 2017

APPROVAL FOR SUCCESSFUL DEFENSE

Masters Candidate, Victoria Bartek, has successfully defended and made the required modifications to the text of the master's thesis for the M.S. during this Spring Semester 2017.

THESIS COMMITTEE
(please sign and date beside your name)

Mentor:
Marianne Lloyd, PhD:________________________________

Committee Member:
Kelly Goedert, PhD:________________________________

Committee Member:
Leamarie Gordon, PhD:________________________________

Acknowledgements

I’d like to thank my Faculty Advisor Dr. Marianne Lloyd, whose guidance and patience has be instrumental at every step of the thesis process. Her support and instruction throughout my time at Seton Hall as a researcher and a graduate student have helped me immensely to grow in knowledge and confidence as not only researcher, but also as a student and overall person.

I’d also like to thank both members of my thesis committee, Dr. Kelly Goedert and Dr. Leamarie Gordon, for their incredible encouragement and insight throughout the past 2 years. The time and knowledge that they invested in directing me through the thesis process have been invaluable.

Lastly, I’d like to thank the members of the Seton Hall Memory Lab, for their time in collecting data and providing suggestions for this study and prior pilot studies. It has been a pleasure to work with this group of intelligent and dedicated people, and I would have never been able to collect all the data on time without their help.

Table of Contents

List of Figures

List of Tables

Abstract

Retrieval enhanced suggestibility (RES) refers to an effect where initial testing of an event leads to better learning of and higher production of misinformation regarding that event. This paper proposes the New Theory of Disuse (Bjork & Bjork, 1992) as a supplement to the retrieval fluency account for RES (Thomas et al., 2010). The amount of interference presented between the misinforming narrative and final test was manipulated in order to investigate how decays in retrieval strength (how easily a memory is recalled) affect misinformation reporting. Results suggested that the learning of interfering information may decrease RES, but that this effect may be contingent on how strongly the original event memory was stored (quantified as performance on initial test). This is in line with New Theory of Disuse's predictions, which suggest that degree of retrieval strength decay with new learning may be determined by how strongly a memory trace is stored.

Introduction

Loftus is credited with first introducing the world to the misinformation effect. By presenting participants with slides depicting a car accident, then later giving these participants misinformation on this event, Loftus et al. (1978) illustrated that recall for a witnessed memory can be modified by suggested misleading information. This phenomenon has been most commonly applied to errors in our criminal justice system. In particular, the Innocence Project, an organization dedicated to exonerating the wrongly convicted, estimates that eyewitness misidentification, where a person misremembers or misidentifies details of a person or an event, has played a role in over 70% of discovered wrongful convictions since 1992 (The Innocence Project, 2015). Consequently, applied research has attempted to identify procedures to minimize the misinformation effect, such as decreasing feedback given to eyewitnesses following a report (Wells et al., 1998).

Cognitive psychologists have attempted to understand the misinformation effect by identifying factors that facilitate and attenuate suggestibility. The present work focuses on how initially memory for a witnessed event, prior to providing misinformation (i.e., information that directly contradicts that of the original event), affects susceptibility to misinformation (Chan et al., 2009; Thomas et al., 2010; Chan & Langley, 2011; Chan & Lapaglia, 2013; Gordon & Thomas, 2014; Gordon, Thomas, & Bulevich, 2015). These studies consistently illustrate that initial testing decreases accuracy for misleading (changed) details (i.e., less retrieval from the original event) and increases reporting of details from the second, misinforming narrative (i.e., increased misinformation production) beyond what is seen in the standard misinformation effect procedure. This facilitation of the misinformation effect has been termed retrieval enhanced suggestibility (RES).

Why Does Testing Enhance Misinformation Production?

RES is particularly interesting because it seems to defy past research on the testing effect, which suggests that testing enhances information retention, and therefore should improve rather than deter retrieval performance (Butler, 2010; Butler & Roediger, 2007; Roediger & Karpicke, 2006). While some have theorized that RES occurs due to affected retention of the original event memory's content (see Chan & LaPaglia [2013] for research on the reconsolidation theory of RES), most research suggests a retrieval fluency account: where original event details are not overwritten by misinformation, but are more difficult to retrieve than the misinformation. This suggests that the decreases in accurate reporting and increases in misinformation production that characterize RES are potentially reversible. The retrieval fluency account is supported by studies illustrating that when prompted at final test (i.e., by changes in directions or by being warned that they've been exposed to misinformation in the past) participants are able to engage in effortful retrieval and access original event details quite easily (Thomas et al., 2010; Gordon & Thomas, 2014; Gordon et al., 2015; Gordon, 2015). This data supports the retrieval fluency account, but leaves researchers to determine why testing makes misinformation easier to retrieve than the original event.

Using the New Theory of Disuse to Explain RES

Applying Bjork and Bjork's (1992) New Theory of Disuse may help in understanding why and when RES occurs. The New Theory of Disuse suggests that memory can be dissociated into two separate strengths: storage strength and retrieval strength.

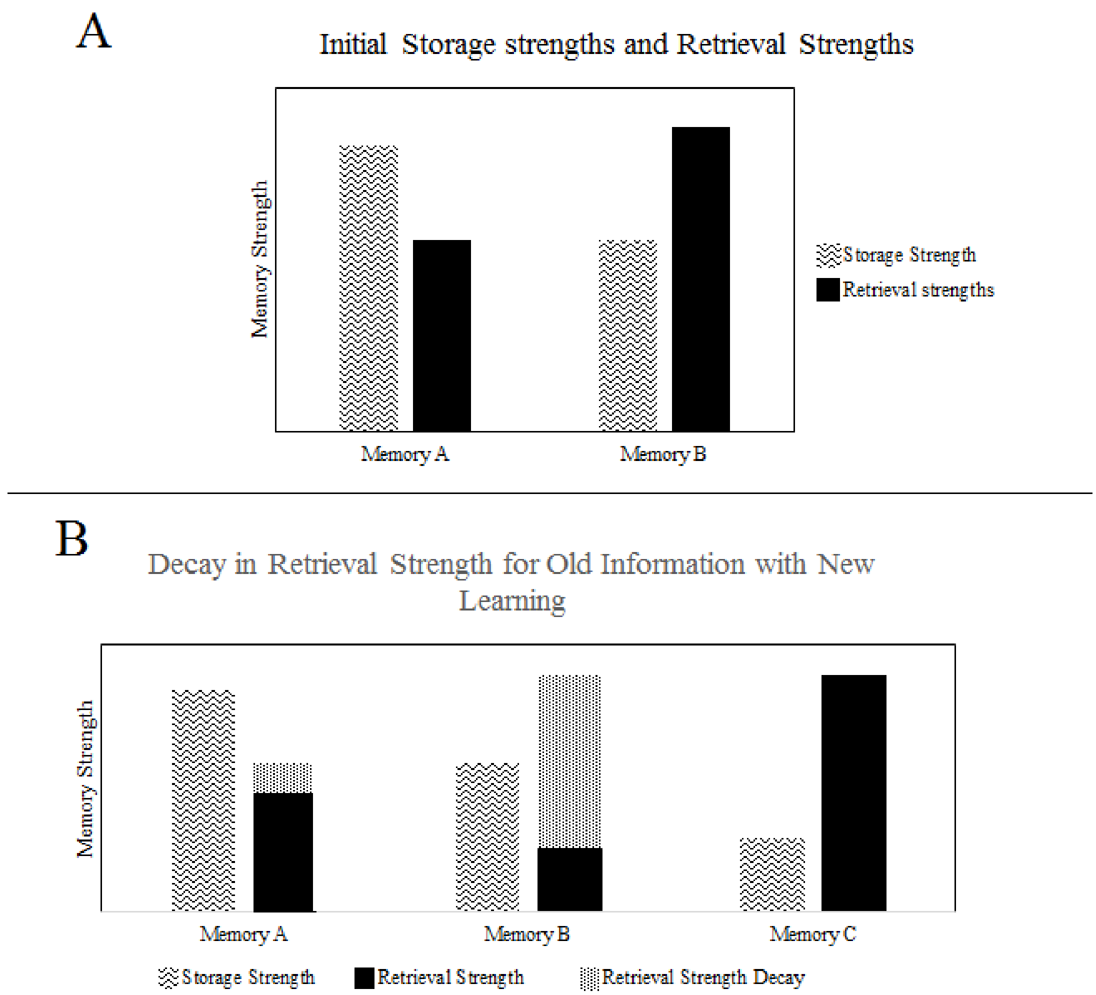

Figure 1. New Theory of Disuse, and the Effect of New Learning on Retrievability (Bjork & Bjork, 1992)

The storage strength of a memory trace is a measure of how well an item is learned. It is increased by opportunities to study or retrieve the item, meaning that items that are used or viewed more often will be more strongly stored. Further, the learning of new items does not affect the storage strength of old items.

Conversely, the retrieval strength of a memory trace determines how easily that item is accessed via retrieval – where items with higher retrieval strength are more easily recalled. The retrieval strength of an item is affected by new learning, where the addition of more items decreases the retrieval strength of already encoded items. Critically, the degree of retrieval strength decay with new learning is dependent on the memory trace's storage strength. While weakly stored items will decay in retrievability dramatically when new information is learned, items that are strongly stored will be less affected by the learning of new information. For example, a memory with a strong storage strength (e.g., an old home phone number; depicted in Figure 1 as Memory A) may be harder to retrieve than a newly learned memory with a weak storage strength (e.g., a friend's new phone number; Memory B) simply because learning of Memory B was more recent. However, if another new memory is encoded (e.g. a string of random numbers; Memory C), Memory B will decrease in retrieval strength more than Memory A due to the influence of a weaker storage strength for Memory B.

The New Theory of Disuse and the testing effect suggest that testing will enhance the retrieval and storage strength of the tested information – therefore predicting improved original event memory (Bjork & Bjork, 1992; Roediger & Karpicke, 2006; Butler & Roediger, 2007; Butler, 2010). However, RES research consistently illustrates the opposite, where testing of the original event increases retrievability of misinformation. Why is there this contradiction between testing theory and the RES phenomenon? The current research proposes two characteristics of the RES procedure as potential explanations: initial test performance and the recency of misinformation at final test.

Initial Test performance and RES

The beneficial effects of initial testing are contingent on whether the test was successful. While the successful retrieval of information may enhance its storage strength (subsequently slowing retrieval strength decay), failed retrieval without feedback does not enhance retention

(Storm et al., 2014; Rowland, 2014). Therefore, in order for testing to enhance original event memory in RES, it needs to be successful. RES studies have reported initial test performance averages around 50% to 60% (e.g.: Gordon et al., 2015: 57%; Chan & Langley, 2011: 58%; unpublished data from Seton Hall University participants: 47%). This relatively low success rate likely provides minimal test enhanced retention of the original narrative.

Further, Gordon and colleagues have suggested that initial testing increases attention to the *misinforming* narrative, as illustrated by increased reading times for test critical details. Their results suggest that this allocation of attention may in fact enhance learning of misinformation, increasing misinformation report at final test (Gordon & Thomas, 2014; Gordon et al., 2015; Gordon, 2015). Therefore, the unsuccessful initial testing often seen in RES may leave the storage strength and retrieval strength of the original event unaffected, while enhancing encoding (and subsequently storage strength and retrieval strength) of misinformation.

Recency of Misinformation and RES

Even when testing is successful its beneficial effects on retention may be minimized by the traditionally short and controlled duration between the misinforming narrative and final test in typical RES procedures. The New Theory of Disuse suggests that the retrieval strength of information decays as new items are learned. The speed of this decay is dependent on how strongly a memory is stored (i.e., more strongly stored items decrease in retrievability more slowly than weakly stored items). However, if additional encoding does not occur prior to final test, retrieval strength for the most recently learned information will remain high, even if the storage strength is much lower than that of older information (see Memory C in Figure 1B; although it is a weakly stored memory its retrievability is high because of its recency). Testing effect studies have illustrated the importance of retention period by showing that the benefit of testing decreases and sometimes even disappears with its reduction (Wheeler, Ewers, & Buonanno, 2003; Halamish & Bjork, 2011). In these studies testing is still presumed to increase the storage strength of information, but because there is no opportunity for retrieval strength decay to occur, the beneficial effects of testing (via improvements in storage strength) on retrieval may not be evident when compared to the performances of participants who passively restudied (Bjork & Bjork, 1992). In the case of RES, while the tested information (i.e., the original event) may be more strongly stored than the misinforming narrative, misinformation may still be more retrievable at final test simply because no additional learning occurred to decrease the retrievability of the misinformation.

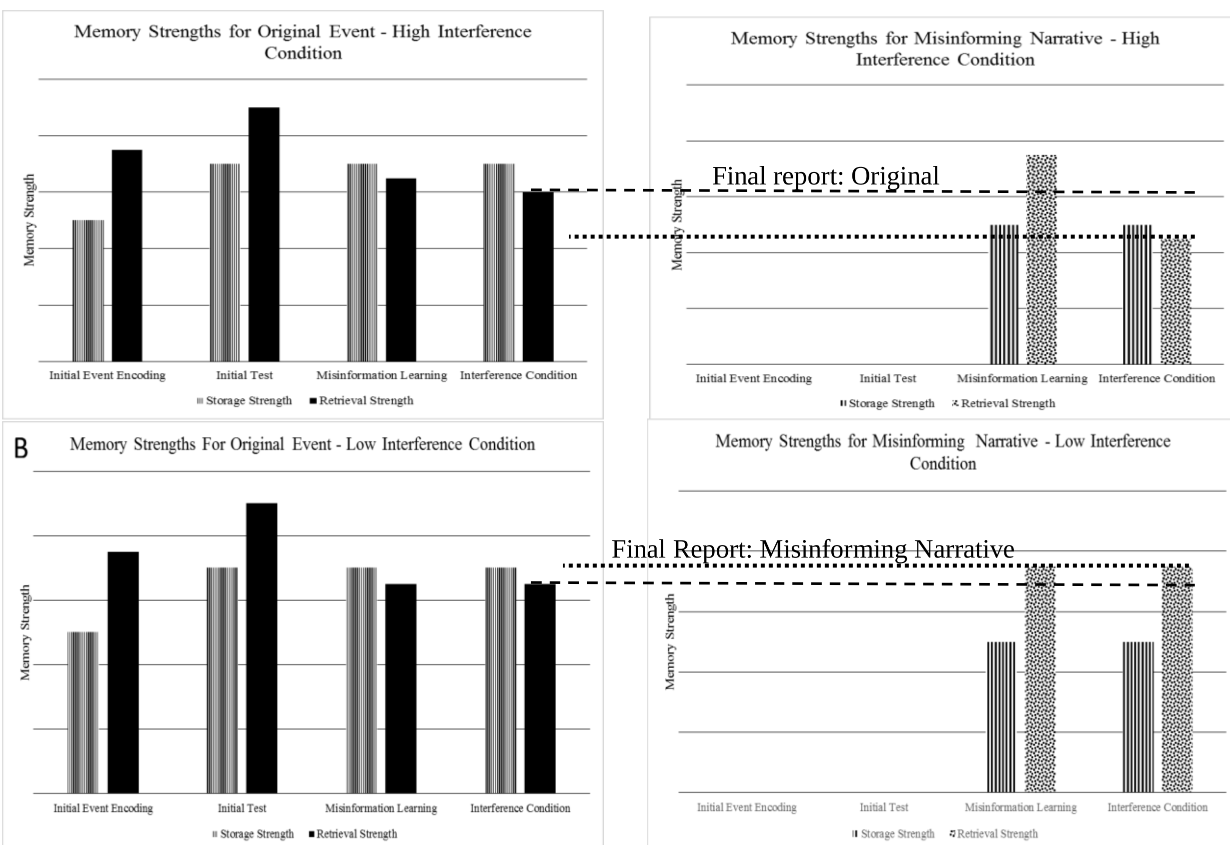

Figure 2. Graph of Hypothesized Storage and Retrieval Strength for Low and High Interference Conditions.

Current Research

The current study investigates how the introduction of an additional encoding period between the misinforming narrative and the final test (i.e., learning an interfering narrative) will affect final test performance and misinformation production. Encoding additional information that is similar to the original and misinforming narratives provides an opportunity for misinformation to decay in retrieval strength prior to final test, an effect that is not expected when participants encode unrelated (i.e., noninterfering) information (see Figure 2 for the hypothesized effect of interference on retrieval strengths at final test). Importantly, this interference manipulation is predicted to decrease RES only for participants who perform well on the initial test. If testing enhances the original event's storage strength so that it's greater than the storage strength for misinformation, it is expected that new high interference learning will affect retrievability of misinformation more than it affects the original event. On the other hand, poor performance on the initial test should minimally affect the storage strength of the original event so that it is equal to, or possibly weaker, than that of the misinformation. Therefore the rate retrieval strength decay would be relatively equal for the two memories and new learning would not affect RES. In the low interference condition initial test performance is not expected to affect final test RES. If there is no opportunity for the misinformation to decay in retrieval strength, it will be more easily recalled than the original event at final test regardless of differences in storage strengths.

Method

Participants

187 Seton Hall University students participated for course credit for a non-descriptive "Memory for Text" experiment.

Materials

Materials were identical to those used in studies from Gordon and colleagues (2014; 2015; 2015). These included a narrative version of the first episode of 24 and a misleading version of this narrative in which 8 test critical details were omitted (“control” details) and 8 details were changed (“misleading”). Each details’ status (consistent, control, or misleading) was counterbalanced across participants. A cued recall test made up of 24 questions was also used. This experiment implemented a text-based original event, while other misinformation effect and RES studies have commonly used an initial video or picture based event (Gordon & Shapiro, 2012; Chan & Langley, 2011; Chan & LaPaglia, 2013, Gordon & Thomas, 2014; Gordon et al., 2015). Unpublished pilot data collected from Seton Hall students suggests that RES persists in this modified procedure, and the modification was made in order to match modality between the original event, misinforming narrative, and the novel text based interference manipulation. This was in response to past research suggesting that matches in modality between test and information may affect final test reporting (Campbell et al., 2007; Abeles & Morton, 1999; Pezdak & Greene, 1993).

Participants were also exposed to either a high interference or a low interference narrative prior to final test. Both narratives were 67 sentences long. The high interference text discussed the aftermath of a break in and contained details similar to those seen in the 24 narrative (e.g. in the original narrative Kim lives on 10th street, in this narrative Frank Gaines lives at 2170 Powell street). The low interference text contained details unrelated to the original and misleading narratives (i.e., discussed honey bees and the consequences of colony collapse disorder).

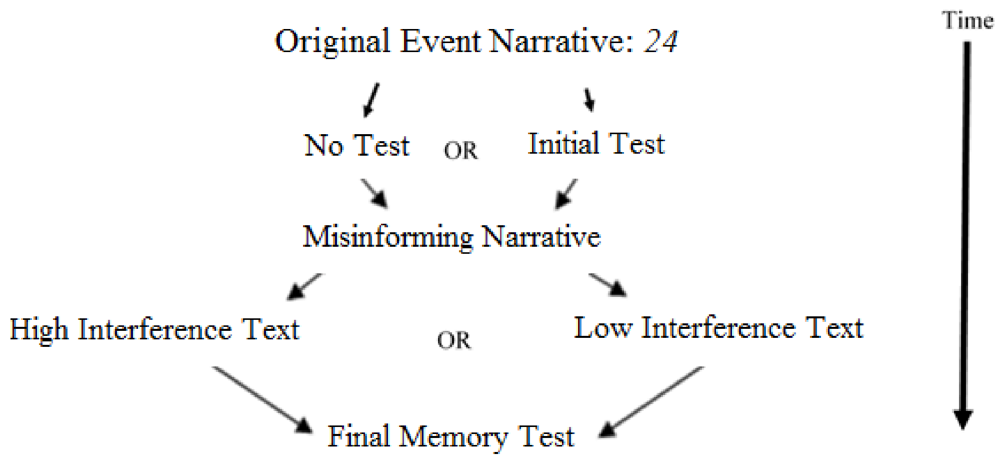

Figure 3. Current Procedure

Procedure

Figure 3 illustrates the procedure. Procedures were similar to those in previous RES studies (Gordon et al., 2015; Gordon & Thomas, 2014; Gordon, 2015), with the exception of using a narrative version of the original event (the first episode of the show "*24"*) and the addition of an interference condition.

After completing an informed consent, all participants were presented a written narrative of the first episode of the show "*24*" sentence by sentence, where they used a space bar press to move from one sentence to the next. Participants were notified that they would be tested later on this narrative. Participants were randomly assigned to either the tested or standard condition. Tested participants performed a 24 question, 6 minute cued retrieval test on the encoded narrative, where they had 20 seconds to answer each question. Standard participants performed a sudoku puzzle for 6 minutes as a filler task. Both sets of participants were then presented the misleading narrative sentence by sentence.

The interference condition was presented to participants immediately following misinformation reading. Participants were randomly assigned to either the high interference or low interference condition. In both cases participants read a 67 sentence narrative unrelated to the original event. The high interference narrative contained details similar to the original event and misinforming narratives presented earlier, while the low interference narrative contained dissimilar information. Finally, all participants performed a cued recall test consisting of the same 24 questions presented to half of the participants earlier, and were asked to report details from the original event.

Results and Discussion

Initial Test Performance

Of the 187 participants in the current experiment 93 were in the tested condition. These participants' scores on the initial test ranged from .08 to .96, reporting a mean proportion of .49 (*SD*=.22) answers successfully.

Effects of Initial Testing and Interference on Final Test Accuracy

Participants' accuracy scores at final test were collected for consistent, control, and misleading details. Overall accuracy at initial test and proportion of misinformation reported at final test were also recorded. Tables 1 and 2 provide mean accuracy rates and mean misinformation production separated by initial test condition and interference group. Results concerning misleading item accuracy and misinformation production are reported first, as these are most central to the current prediction.

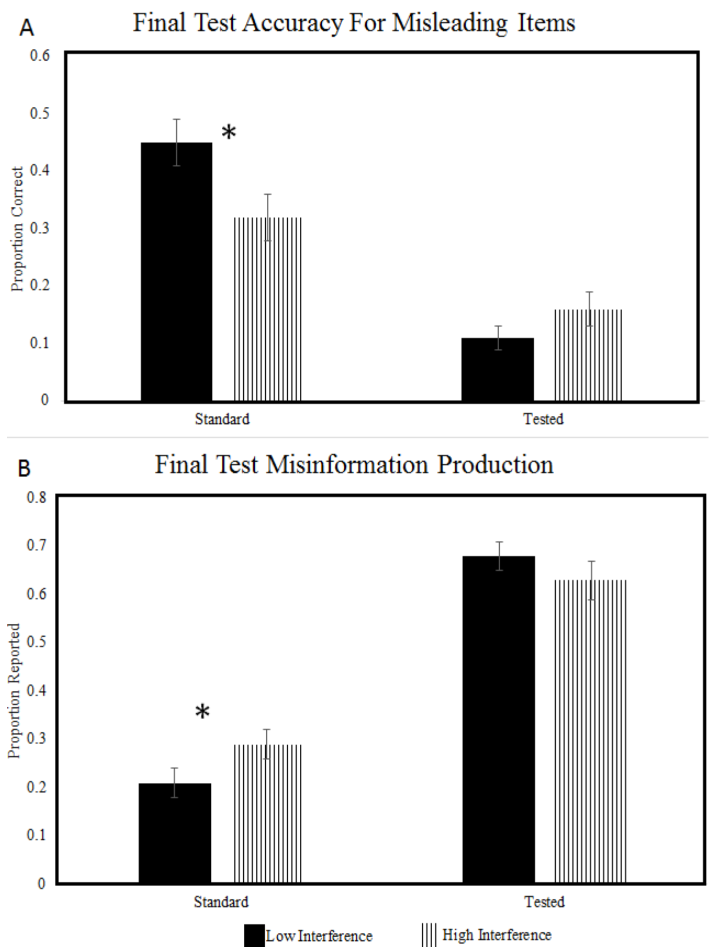

Figure 4. Differential Effects of Interference on Final Test Accuracy and Misinformation Production

Past RES studies illustrate decreases in accuracy on misleading items along with increases in misinformation production with initial testing (Chan et al., 2009; Thomas et al., 2010; Chan & Lapaglia, 2013; Gordon et al., 2014; Gordon & Thomas, 2015). The current analyses were conducted to observe this effect of initial testing and to determine whether encoding interference prior to final testing affects final test performance. Initial analyses used a 2 (initial test condition: tested or standard) x 2 (interference condition: high interference or low interference) between subjects ANOVA to investigate accuracy for misleading items, as well as a

2 (initial test condition: tested or standard) x 2 (interference condition: high interference or low interference) between subjects ANOVA for misinformation production.

Initial Testing and Interference Affect Suggestibility to Misinformation. In line with past RES studies, initially tested participants were less accurate on misleading items (*M*=.14, *SD*=.21) than participants in the standard (*M*=.38, *SD*=.24) condition, $F(1,182)=56.22$, $p<.001$, $\eta^2_p=.24$. They were also more suggestible to misinformation at final test (Tested: *M*=.65, *SD*=.26; Standard: *M*=.45, *SD*=.24), suggesting that initially tested participants were not only failing to retrieve the correct answers but were answering these questions with misinformation, $F(1,182)=144.69$, $p<.001$, , $\eta^2_p=.44$.

Further, high interference learning increased suggestibility to misinformation for participants in the standard condition. As illustrated in Figure 4, high interference decreased misleading item accuracy (*M*=.31, *SD*=.21) and increased misinformation production (*M*=.30, *SD*=.20) at final test as compared to the low interference group (misleading item accuracy: *M*=.45, *SD*=.24; misinformation production: *M*=.21, *SD*=.18), $F(182)=8.35$, $p=.004$, $\eta^2_p=.05$; $F(1,182)=4.22$, $p=.04$, $\eta^2_p=.02$. However, high interference learning did not affect suggestibility in tested participants, *p*'s >.05. This suggested that while interference may make retrieval for the original event more difficult, initial testing may protect against this effect. Across test conditions interference did not significantly affect final test accuracy or misinformation production, *p*'s >.05.

Table 1.

Final Test Accuracy for Consistent, Control, and Misleading items organized by Interference and Initial Test Conditions

Initial Test Condition	Low Interference	High Interference
Not Tested	.21 (.18)	.30 (.20)
Tested	.67 (.23)	.63 (.28)

Note. Standard deviations are included in parentheses.

Table 2.

Average misinformation production organized by test condition and interference condition.

Interference Condition	Initial Test Condition	Consistent	Control	Misleading
Low Interference				
	Standard	.75 (.22)	.47 (.24)	.45 (.24)
	Tested	.80 (.20)	.28 (.25)	.11 (.16)
High Interference				
	Standard	.63 (.27)	.36 (.22)	.31 (.21)
	Tested	.82 (.17)	.35 (.31)	.17 (.25)

Note. Standard Deviations are given in parentheses

Initial Test Performance and Suggestibility to Misinformation. The initial analyses suggest that initial testing and interference interact to affect final test performance – where initial testing protects against the increased suggestibility to misinformation caused by high interference. However, in order to address the hypothesis (that high interference learning will decrease suggestibility to misinformation contingent on the storage strength of the original event memory) it is necessary to take initial test performance into account. In order to do so, separate simple linear regressions on were completed on tested participants' data, where initial test performance was used to predict RES in low and high interference conditions. The New Theory of Disuse predicts that original event memory will be more strongly stored than the misleading event if testing is successful (due to the testing effect), and that providing an opportunity for retrieval strength decay (i.e., high interference) will decrease retrievability of the misleading narrative and increase reporting from the original event. Therefore, improvements in initial test accuracy should predict higher accuracy on misleading items and decreases in misinformation production at final test for the high interference condition. Data from these analyses are reported in table 3.

As predicted, high interference participants were more accurate on misleading items at final test (i.e., reported more from the original event narrative) as success on initial test performance increased. β=.32, $F(1,46)=5.22$, $p=.03$, $R^2=.10$. This is illustrated in figure 5. In contrast, the low interference condition showed no relationship between initial test performance and accuracy on misleading items ($p= .56$), suggesting that while improvements in initial test performance increased storage strength for the original event, interference had to be applied in order to observe an effect on retrieval.

Table 3.

Predictive ability of initial test performance for RES in Low and High Interference Conditions

Criterion	F	Constant	R^2	b	β	p
Interference Condition						
Low Interference						
Accuracy on Misleading Items	.35	.08	.008	.07	.09	.56
Misinformation Production	3.52	.53	.08	.32	.28	.07
High Interference						
Accuracy on Misleading Items	5.22*	-.19	.10	.35	.32	.03
Misinformation Production	1.88	.50	.04	.24	.20	.18

The current hypothesis also predicted that the high interference condition would report less misinformation as initial test performance improved. This result was not observed, p=.18. For the low interference group improvements in initial test performance marginally predicted higher rates of misinformation (β=.28, F(1,43) = 3.52, p=.07, R^2=.05), suggesting that better initial test performance actually facilitated RES in the low interference condition. However, this relationship became nonsignificant after participants who received a perfect "RES score" at final test (i.e. reported all of the misinformation answers at final test and got no misleading items correct; 13 participants) were removed (p=.81). It is possible that these participants purposely reported from the misleading narrative, due to a misunderstanding of the directions. This suggests that the relationship between misinformation production and initial test performance may have been due to a confusion rather than an effect of low interference. No other results were affected by the removal of these participants.

Effects of Initial Testing and Interference on Final Test Accuracy. While the current predictions focused primarily on the effects of initial testing and interference on RES, changes in consistent and control item accuracy were also of interest. In past RES studies, initial testing has increased performance on consistent items, an effect suggested to be due to testing increasing attention to test critical details (i.e., the attention allocation hypothesis; Gordon & Thomas, 2014; Gordon, 2015; Gordon et al., 2015). Further, while control item accuracy is inconsistently benefited by initial testing in RES studies, it provides an example of typical test enhanced retention (Gordon & Thomas, 2014; Gordon et al., 2015). A 2 (Tested or Standard) by 2 (Low Interference or High Interference) by 3 (Consistent, Control, or Misleading) mixed measures ANOVA (conducted in line with past statistical procedures for RES; Gordon et al., 2015; Gordon & Thomas, 2015; Thomas et al., 2010; Chan & Langley, 2011) revealed the effects of initial testing and interference on overall performance as well as on accuracy on specific item types (consistent, control, or misleading). While no predictions were made regarding the effect of interference on accuracy, initial testing was expected to increase performance on consistent items. Further, it was expected that, overall, participants would perform better on consistent items than on control or misleading items.

Participants' success on test items was affected by question type, $F(2,364)=372.58$, $p<.001$, $\eta^2_p=.67$. Performance was more accurate for consistent items ($M=.75$, $SD=.23$) than for control ($M=.36$, $SD=.27$) or misleading ($M=.26$, $SD=.26$; $t(185)=18.83$, $p<.001$, $d=1.55$, $\alpha = .017$; $t(185)=20.93$, $p<.001$, $d=2.03$, $\alpha = .017$) and more accurate for control items than for misleading items, $t(185) = 5.36$, $p< .001$, $d=.39$, $\alpha = .017$. These findings match those in past RES studies and make intuitive sense; details that were presented twice were more accurately remembered than those that were only presented once or those that were changed.

Overall, participants who were initially tested performed worse on the final test (*M*=.42, *SD*=.17) than those in the standard (*M*=.49, *SD*=.20) condition, F(1,182)= 7.92, p = .005, η^2_p = .04. This was due to initial testing negatively affecting performance on control (Tested: *M*=.31, *SD*=.29; Standard: *M*=.41, *SD*=.23; t(177.30)=2.67, p=.008, d=.39) and misleading (Tested: *M*=.14,*SD*=.21; Standard: *M*=.38, *SD*=.24) items, t(181.507)=7.29, p<.001, d=1.07. Meanwhile, as seen in past studies, testing enhanced performance on consistent items (Tested: *M*=.81, *SD*=.18; Standard: *M*=.69, *SD*=.25; t(167.55)=3.86, p<.001, d=.57) illustrating a differential effect of testing on item types, F(2,364)= 48.15, p<.001, η^2_p=.21. This suggests that while testing may hurt retention of original event details (i.e., misleading or control), it enhances relearning of consistent details during encoding of the misinforming narrative.

While there was no overall effect of interference on final test performance (p=.16), interference did differentially affect test conditions, F(1,182)= 10.13, p=.002, η^2_p=.05. Standard condition participants who encoded high interference information performed worse overall at final test (*M*=.43, *SD*=.19) than those in the low interference condition (*M*=.55, *SD*=.19), t(91)=3.10, p=.003, d=.64. However, initial testing eliminated this effect, as tested participants' overall final test performance was similar between low interference (*M*=.40, *SD*=.14) and high interference (*M*=.44, *SD*=.19) conditions, p=.19. There was no significant interaction between interference, test condition, and question type, p=.26.

Therefore, while initial testing affected final test accuracy overall by decreasing control and misleading item accuracy, it also protected against the decrements in performance caused by interference. The new theory of disuse may provide an explanation for this interaction between initial testing and interference. Initial testing may increase the storage strength for the original event so that its retrieval strength is less affected by this new interfering learning.

Initial Test Performance Predicts Final Test Accuracy. Arguably the most counterintuitive aspect of the current results was the detrimental effect of initial testing on control items accuracy – a finding that directly contradicts the testing effect. One possible explanation is that initial test performance was not taken into account for these analyses. Since initial test performance was low (M=.49, SD=.22), many participants performance may not have been successful enough to provide the typical beneficial effects of testing. Linear regressions using initial test performance to predict final test accuracy for control items support this suggestion for both high interference and low interference conditions. Increases in initial test performance predicted success on control items at final test, low interference: β=.36, $F(1,43)=6.23$, $p=.02$, $R^2=.13$; high interference: β=.38, $F(1,46)=7.79$, $p=.008$, $R^2=.15$. This is in line with past research on the testing effects suggesting that in order for testing to enhance retention it must be successful (Rowland, 2014). Participants with high initial test performances also were more successful on consistent items for both high and low interference conditions, β=.65, $F(1, 43)=30.55$, $p<.001$, $R^2=.42$; β=.56, $F(1,46)=20.68$, $p<.001$, $R^2=.31$. This was not surprising because these details did not differ between the original event and the misleading event. Participants who encoded this information well at first test likely retained the information (i.e., test enhanced retention) and also encoded the information more effectively during its second presentation in the misinforming narrative (i.e., test enhanced learning; Roediger & Karpicke, 2006).

Reading Time Analysis. In the current study, participants' accuracy on misleading items was significantly predicted by initial test performance in the high interference group. This supports a retrieval fluency account of RES by illustrating that the original event memory is available and can become more accessible via manipulations of retrieval fluency (or, as put by the New Theory of Disuse, retrieval strength).

In line with past RES studies supporting the retrieval fluency account, reading time analyses were completed on consistent, control, and misleading sentences using a 3 (Item type: consistent, control, misleading) x 2 (Initial Test Condition: tested, standard) mixed measures ANOVA on participants' median reading times. Past research supporting the retrieval fluency account suggests that increased attention to consistent and misleading details after testing may facilitate misinformation reporting (Gordon et al., 2014; Gordon & Thomas, 2015; Gordon, 2015). Therefore it was hypothesized that tested participants in the current study would spend more time reading consistent and misleading test critical details than the standard condition. The effect of interference on reading times was not included in the analysis because the manipulation was applied directly after all reading times of interest were recorded.

Consistent with Gordon et al. (2015), reading times faster than 300 ms were excluded. Mean reading times and standard error values organized by test condition and item type are presented in Table 4.

Participants spent different amounts of time with items types ($F(2,370)= 24.08$, $p <.001$, $\eta^2_p=.39$), attending to misleading items for longer amounts of time than consistent or control items $t(186) = 7.57$, $p <.001$, $d=.36$;$t(186)=13.84$, $p < .001$, $d=.87$, $\alpha s=.017$. Participants also attended to consistent items for longer amounts of time than control items, $t(186)=8.45$, $p<.001$, $d=.50$, $\alpha=.017$. This suggests that overall participants were paying more attention to details critical to answering the initial and final test questions.

Initial testing increased the amount of time participants spent reading the misleading narrative, $F(2,370) = 119.56$, $p < .001$, $\eta^2_p=.39$. This pattern was contingent on item type ($F(2,370)=7.18$, $p = .001$, $\eta^2_p=.04$), where tested participants spent more time reading consistent information ($t(185)=2.93$, $p = .004$, $d=.43$) and misleading information ($t(185)=3.01$, $p=.003$, $d=.44$) than standard group participants. There were no significant differences in reading times for control items, $p=.32$.

Table 4

Mean reading times for Consistent, Control, and Misleading Sentences in the misinforming narrative.

Condition	Consistent	Control	Misleading
Standard	3680.71 (*163.60*)	3239.51 (*132.31*)	4238.15 (*166.79*)
Tested	4344.82 (*156.33*)	3405.05 (*102.37*)	4992.35 (*186.48*)

Note. Standard error values are given in parentheses.

These results are in line with those illustrated in Gordon and Thomas (2014) and Gordon et al., (2015), where participants who were initially tested spent significantly more time attending to consistent and misleading test critical details than participants in the standard condition. Further, mean reading times and standard error values are comparable between this study and results from Gordon et al. (2015), suggesting that encoding the original event in text form did not affect later attention to the misinforming narrative.

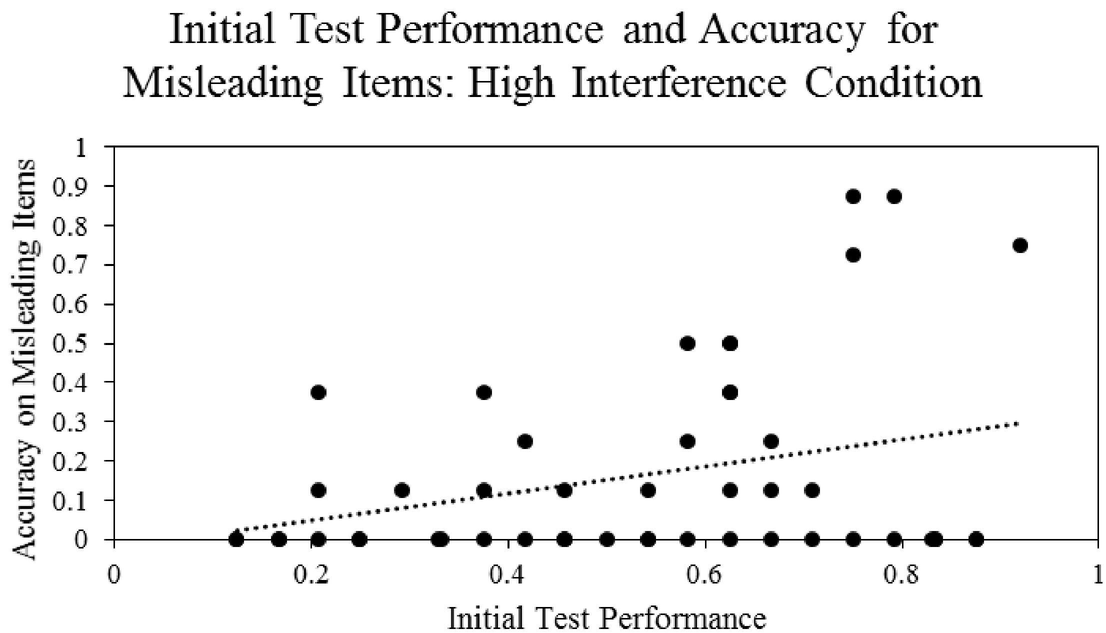

Figure 5. Increases in Initial Test Performance Predicted Higher Accuracy for Misleading Items in the High Interference Condition (β=.32).

General Discussion

The current study framed RES within the New Theory of Disuse in order to observe how the differences in storage and retrieval for both the original event and misleading narrative affected final test RES. The New Theory of Disuse suggests that memory traces have separate storage and retrieval strengths, and that the rate of retrieval strength decay is contingent on the item's storage strength (where items that are more strongly stored decay more slowly). While testing effect studies and the New Theory of Disuse have suggested that initial testing may enhance storage strength for tested items, past RES research consistently illustrates that testing decreases final test accuracy on misleading items and increases misinformation production. The current study illustrated that this effect of testing on RES may be due in part to poor initial test performance (i.e., minimal benefit to storage strength) and to a short, non-interfering interval between misinformation learning and final test (no opportunity for misinformation to decrease in retrieval strength).

The Effect of Interference in RES

The current results illustrated that learning new interfering information prior to final test differentially affected accuracy and misinformation production in initially tested and standard condition participants. While standard participants performed more poorly after high interference encoding, tested participants were seemingly protected from these deleterious effects. Further, analyses focusing solely on initially tested participants illustrated that the RES effect could be significantly predicted by initial test performance in the high interference condition. Increases in initial test performance predicted accuracy on misleading items for the high interference condition (i.e., a decrease in RES). This result persisted after removing participants who may have been confused by the directions. These results are in line with our hypothesis: that the encoding of highly interfering information would decrease the RES effect for participants who performed well on the initial test.

The current results contradict past RES studies that applied extended retention periods, which have not illustrated RES reduction with a retrieval strength manipulation (in this case by providing more opportunities for decay via time; Chan & Lapaglia, 2013; Chan & Langley, 2011). The results of the current study suggest that this persistence of RES in past studies is due in part to low initial test success. While the New Theory of Disuse suggests storage strength increases with successful retrieval, failed retrieval without feedback does not enhance retention (Halaimish & Bjork, 2011; Rowland, 2014). Therefore, it is possible for the original event memory to decrease in retrieval strength more rapidly than the retrieval strength of misinformation, despite initial testing. This was supported by the current study's results, where initial test success predicted reductions in RES at final test within the high interference condition.

Initial test performance also predicted increased misinformation production in the low interference condition. However, because this effect disappeared with the removal of 13 “perfect RES” participants it may have been due to misinterpretation of the directions rather than an effect of retrieval strength. In order to alleviate confusion and to ensure participants are reporting the most fluent answer at final test, a future study may consider asking participants to report the first valid answer that comes to mind. Further, Thomas et al. (2010), illustrated that effortful retrieval at final test took significantly more time than typical reporting. By measuring the amount of time needed for retrieval, future research may be able to determine how interference affects fluency of RES and original event answers.

Importantly, the effects of interference on final test RES more strongly support a retrieval fluency account of RES rather than a reconsolidation account. If, as suggested by the reconsolidation account, RES is in part due to initial testing increasing memory modification, encoding an interfering narrative may be expected to decrease misinformation reporting (due simply to an effect of new learning on memory retrieval) but increases in original event reporting, as seen here, would not be predicted. This ability to retrieve intact original event memories after interference suggests that the original event memory is simply inaccessible to due to the high retrieval strength of misinformation, rather than to a modification of the memory.

Conclusion

Past research on retrieval enhanced suggestibility has illustrated that initial testing increases misinformation production and decreases accuracy for misleading items at final test. The current study replicated the typical RES effect, and approached RES from a novel perspective by manipulating the amount of interference in the typical procedure and considering the relationship between RES and initial learning performance.

The current study supports the current retrieval fluency account of RES by illustrating that the original event is in fact accessible at final test. Further, this accessibility may be facilitated by providing an opportunity for the misinforming narrative to decay in retrieval strength (via encoding of interfering information), an effect that may be contingent on success on the initial test.

Critically, these findings are in line with predictions from the New Theory of Disuse, suggesting that dissociating the effects of testing on retention and on retrievability may provide a better understanding of the mechanisms of RES.

References

Abeles, P., & Morton, J. (1999). Avoiding misinformation: Reinstating target modality. *The Quarterly Journal of Experimental Psychology: Section A*, *52*(3), 581-592.

Bjork, R. A., & Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. *From learning processes to cognitive processes: Essays in honor of William K. Estes*, *2*, 35-67.

Butler, A. C., & Roediger III, H. L. (2007). Testing improves long-term retention in a simulated classroom setting. *European Journal of Cognitive Psychology*, *19*(4-5), 514-527.

Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(5), 1118-1133.

Campbell, J. M., Edwards, M. S., Horswill, M. S., & Helman, S. (2007). Effects of contextual cues in recall and recognition memory: The misinformation effect reconsidered. *British Journal of Psychology*, *98*(3), 485-498.

Chan, J. C., Thomas, A. K., & Bulevich, J. B. (2009). Recalling a Witnessed Event Increases Eyewitness Suggestibility The Reversed Testing Effect. *Psychological Science*, *20*(1), 66-73.

Chan, J. C., & Langley, M. M. (2011). Paradoxical effects of testing: Retrieval enhances both accurate recall and suggestibility in eyewitnesses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(1), 248-255.

Chan, J. C., & LaPaglia, J. A. (2013). Impairing existing declarative memory in humans by disrupting reconsolidation. *PNAS*, *110*(23), 9309-9313.

Gordon, L. T., & Shapiro, A. M. (2012). Priming correct information reduces the misinformation effect. *Memory & Cognition*, *40*(5), 717-726.

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., & 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.

Gordon, L. T. (2015). Retrieval and Attention: Factors that Potentiate Learning and Enhance Eyewitness Suggestibility. *Proquest*

Halamish, V., & Bjork, R. A. (2011). When does testing enhance retention? A distribution-based interpretation of retrieval as a memory modifier. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(4), 801-812.

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), 19-31.

Pezdek, K., & Greene, J. (1993). Testing eyewitness memory: Developing a measure that is more resistant to suggestibility. *Law and Human Behavior*,*17*(3), 361-369.

Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning taking memory tests improves long-term retention. *Psychological science*, *17*(3), 249-255.

Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin*, *140*(6), 1432-1463.

Storm, B. C., Friedman, M. C., Murayama, K., & Bjork, R. A. (2014). On the transfer of prior tests or study events to subsequent study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(1), 115

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.

Wells, G. L., Small, M., Penrod, S., Malpass, R. S., Fulero, S. M., & Brimacombe, C. E. (1998). Eyewitness identification procedures: Recommendations for lineups and photospreads. *Law and Human behavior*,*22*(6), 603-647.

Wheeler, M., Ewers, M., & Buonanno, J. (2003). Different rates of forgetting following study versus test trials. *Memory*, *11*(6), 571-580.