# Warnings Improve Eyewitness Memory and Influence Retrieval Strategies in the Context of Misinformation

Thesis submitted by:

McKinzey G. Torrance

Tufts University

April 2019

Advisors: Elizabeth Race, PhD. and Ayanna Thomas, PhD.

## Author Note

McKinzey G. Torrance, Department of Psychology, Tufts University.

First, I would like to thank the National Science Foundation for providing the funding that made this research possible. Second, I would like to thank Dr. Jess Karanian, Nat Rabb, Jess Jimbo, and Camille Carlisle for all your assistance in data collection and analysis. I would also like to thank Dr. Ayanna Thomas for all your guidance and input that was critical to making this project come together. Most importantly, I would like to fondly thank Dr. Elizabeth Race for being my mentor during this entire process and these past three years. Without your encouragement and advice, this project would not have been possible.

Correspondence concerning this article may be addressed to McKinzey G. Torrance, Department of Psychology, Tufts University, Medford, Massachusetts, 02155. E-mail: mckinzeytorrance@gmail.com

Table of Contents

## Abstract

Eyewitness testimony is often used as evidence for convicting someone of a crime, yet eyewitness memory is not always a perfect recount of events. In fact, prior research has shown that when eyewitnesses encounter misleading information after an event, they are more likely to provide incorrect information about the original event when later asked about the crime. This is called the misinformation effect. Furthermore, more recent research has found that the misinformation effect can be enhanced when eyewitnesses are repeatedly asked about the original crime. Although post-warning (warning after the post-event information) has been shown to reduce these negative effects of misinformation, little research has been done to look into the effect of pre-warning (warning before the post-event information). The current study investigated whether pre-warning and post-warning have similar effects on memory in the context of repeated testing and misinformation, and how these types of warning may affect individuals' memory strategies. We found that both pre-warning and post-warning diminished the misinformation effect by comparable degrees. In addition, subjective reports suggested that both warning types modulate memory strategies during the retrieval period, and that the underlying mechanism of warning is source monitoring.


*Key Words: Eyewitness, Misinformation Effect, Pre-warning, Post-Warning, Source Monitoring*

# Warnings Improve Eyewitness Memory and Influence Retrieval Strategies in the Context of Misinformation

When it comes to convicting someone of a crime, an eyewitness' testimony can mean the difference between being found guilty or innocent. This type of testimony is convincing because it is assumed that eyewitnesses are able to recount what they saw without bias since they were not directly involved in the crime. Unfortunately, eyewitnesses are not immune to lapses in memory or memory degradation. For example, depending on how much time has passed after witnessing a crime, an eyewitness can forget details about what they saw, thus leading them to not give a full picture of what happened when they report their account to police. In addition to forgetting details, it has been shown that if participants encounter information after the event that conflicts with what they saw, that their accuracy for the original event is reduced (Loftus, 2005). For example, an eyewitness could view a news special that contains speculative details about the crime or could receive leading questions from investigators when giving their official statement. Then, when they are later asked about the original crime, they are more likely to provide incorrect information about the original event (Loftus, 2005; Frenda, Nichols, & Loftus, 2011). This trend of misleading post-event information interfering with memory of the original event has been called the misinformation effect. It is operationalized within eyewitness memory studies as reduced original accuracy for the original event (e.g., crime) as well as increased reporting of misleading post-event details.

The misinformation effect presents a clear danger to our justice system because it suggests that eyewitnesses could give inaccurate testimony, thus increasing the likelihood of convicting an innocent person. Indeed, data from the Innocence Project indicates that approximately 70% of individuals they exonerated of crimes were wrongly convicted because of faulty eyewitness

reports (Innocence Project, 2016). This has led researchers to investigate the origins of the misinformation effect with regard to eyewitnesses in an effort to find a solution. After years of research, a scientific consensus has been established that eyewitness accuracy can be largely affected by introducing misinformation after viewing a crime (Loftus, 2005). While this research was important for establishing the presence of the misinformation effect, the majority of prior research has only investigated the misinformation effect within one specific type of testing paradigm: the single test paradigm. What this means is that participants in misinformation studies have only one instance in which they report their memory of the original event *after* they have encountered misinformation (i.e., Witnessed Event -> Post-Event Information-> Memory Retrieval; Lindsay, Allen, Chan, & Dahl, 2004; Loftus, Miller, & Burns, 1978). Recently, this three-stage procedure has been seen as constraining because it does not account for participants recounting or revisiting their memory of the crime before encountering misinformation. In the real world, an eyewitness may recount what they saw to a 911 operator in order to report the crime, or simply think back to the crime they saw when going about their day. Such repeated instances of memory retrieval mean eyewitnesses are engaging in memory retrieval before encountering any type of post-event information. This likely memory process highlights the need to adapt the methodology used to study eyewitness memory within lab settings, a need that is made even more pressing given prior research demonstrating that repeated retrieval enhances susceptibility to misinformation.

Chan, Thomas, and Bulevich (2009) created a repeated-retrieval misinformation paradigm where participants performed a memory test before *and* after encountering misinformation and found that susceptibility to misinformation increases in a repeated-retrieval misinformation paradigm. In this study, an initial test of memory occurred before mock eyewitnesses

encountered any post-event information (i.e. Witnessed Event-> Memory Retrieval -> Post-Event Information-> Final Memory Retrieval; Chan, Thomas, & Bulevich, 2009). The authors originally hypothesized that the initial memory retrieval period would act as a consolidation event and lead participants to have a *more* accurate memory of the original crime during the final memory retrieval. Surprisingly, the authors found that the initial test of memory made participants more susceptible to misinformation provided in an audio narrative. Participants who were given an initial test of memory scored significantly lower on the final memory test compared to participants who had no initial test and were ~20% more likely to produce the misinformation from the post-event information (audio narrative) (Chan et al., 2009). This effect was named the Retrieval Enhanced Suggestibility (RES) effect (Chan et al., 2009). The finding of the RES effect demonstrates the dangerous effects of participants retrieving the original memory before encountering misinformation and highlights the importance of investigating misinformation in the context of repeated retrieval.

However, there is not a consensus on how the RES effect is modulating memory. One possibility is that the RES effect reflects memory reconsolidation (Chan et al., 2009). According to this view, initial testing reactivates the original memory and puts it into a labile state that requires reconsolidation, where the original memory is more susceptible to interference. Thus, when a participant encounters misinformation which closely parallels the original memory, this misinformation becomes incorporated into the original memory. While Chan et al. (2009) claimed their findings were tentative evidence for a reconsolidation process in human episodic memory, they also acknowledged that there could be other possible explanations of RES effect. A second possibility is that the RES effect reflects retrieval fluency. According to the retrieval fluency hypothesis, first proposed by Thomas, Bulevich, and Chan (2010), participants better

learn details in post-event information that were asked about during initial testing. This enhanced encoding creates a new memory of the misleading details and increases ease of retrieval. Thus, when participants are asked about these same details in the final memory test, the more recently encoded memory is easier to retrieve and results in the increased susceptibility to misinformation (i.e. RES). In support of this hypothesis, Thomas et al. (2010) found that subjects who were not warned and in a repeated testing condition responded with the most fluently retrieved item (Thomas et al., 2010). They posited that this finding, compared to the higher accuracy of warned participants who also had slower retrieval times, supported the retrieval fluency hypothesis because it showed that when participants took more time to remember the original memory, they were able to access it (Thomas et al., 2010). This finding not only suggested that accessing the original memory is possible after encountering misinformation, but that these memories are distinct since it took varying levels of time to access them. While there is strong evidence in support of the retrieval fluency hypothesis, it is important to note that both theories could be valid, just within different paradigms or at different times within an experiment.

Regardless of the mechanism by which the RES effect modulates memory, we still must consider how might we mitigate susceptibility to misinformation in the more naturalistic context of repeated retrieval. One possibility would be to try to eliminate exposure to misinformation. Unfortunately, while eliminating exposure to misinformation would be the strongest way to ensure it does not affect eyewitness memory, this is an impossible task. In the above example about leading questions, it is often not a conscious act for investigators to lead witnesses, and even in the cases where it is, it would not be feasible to monitor every interaction an investigator had with witnesses. In addition, in the case of eyewitnesses hearing about the crime they saw on the news, it is not realistic to quarantine witnesses from all outside sources of information on the

bases that they *could* encounter misinformation before they make their statement to police. The inability to eliminate misinformation in the environment has led researchers to look for methods to counter the misinformation effect knowing that exposure to misinformation is likely unavoidable.

As described above, one of the main ways that researchers have investigated protecting eyewitnesses from misinformation is through warnings, specifically post-warnings (Thomas et al., 2010). Post-warning is when participants receive a warning about the veracity of the post-event information *after* they have encountered it. Whether dealing with a single test paradigm or a two-test paradigm, the post-warning is placed between Post-Event Information and Memory Retrieval (e.g. Witnessed Event->Post-Event Information-> Warning -> Memory Retrieval; Witnessed Event-> Initial Memory Retrieval -> Post-Event Information-> Warning -> Final Memory Retrieval, respectively). In single test paradigms, post-warning has been shown to significantly decrease the rate of memory errors for questions that were related to the encountered misinformation (Blank & Launay, 2014; Loftus, 2005), reducing the misinformation effect by almost fifty percent in some cases (Blank & Launay, 2014). Post-warning has not only shown to be effective within the single test paradigm, but within repeated testing paradigms as well. Thomas et al. (2010) added a post-warning to their procedure (i.e. Witnessed Event-> Initial Memory Retrieval -> Post-Event Information-> Warning -> Final Memory Retrieval; Thomas et al., 2010) and found that when participants who received an initial test of memory also received a post-warning, the RES effect was eliminated (Thomas et al., 2010). These results suggest that warning consistently reduces the negative effects of misinformation and can be used as a potential tool for enhancing eyewitness memory.

In order to understand how post-warning is effective, it is necessary to investigate its underlying mechanism. Prior research has suggested that misinformation, specifically in regard to eyewitnesses, lends itself nicely to the source-monitoring framework (Johnson, Hashtroudi, & Lindsay, 1993). What this framework suggests is that the misinformation effect is a retrieval error and that participants are referencing the wrong memory when answering final memory tests. Interestingly, past research has suggested just that. Gordon and Thomas (2014) investigated what happened when participants were able to write down two responses to a question, instead of just being forced to choose one during a final memory test. These authors found that participants who were given initial testing compared to participants who received no initial testing had improved access to both original event details and post-event information on misleading trails. This study suggest that participants who receive initial memory tests have two separate memories that they are able to access. Importantly, this finding suggests that participants likely have two memory traces to access and that source monitoring is possible.

In addition to this prior research suggesting the presence of multiple memory traces with different sources within people who had initial testing, the source monitoring framework suggests that manipulations that enhance source memory should reduce the effects of misinformation (Johnson, et al., 1993). Thus, given previously discussed findings about how post-warning lowers the RES effect and the misinformation effect (Thomas et al., 2010; Blank & Launay, 2014), we can explore the possibility that source monitoring represents the underlying mechanism by which post-warning protects memory from misinformation. Indeed, Blank and Launay (2014) speculate that source monitoring could be an underlying mechanism given their own findings. However, more research is required to draw more concrete connections between post-warning and source monitoring and to better understand the mechanisms by which post-

warning affects eyewitness memory. Investigating the mechanisms of warning also raises the question of whether the timing of warnings affects memory performance. More specifically, does providing a warning *before* the post-event information (pre-warning) modulate different aspects of memory processing compared to post-warnings? If so, are pre-warnings potentially more effective than post-warnings at reducing the misinformation effect?

Pre-warnings have already been shown to be more effective that post-warnings in preventing false memory formation in other types of experimental paradigms. In a study by Gallo, Roediger, and McDermott (2001), participants were presented with several word lists, where each word list was comprised of words that were semantically related, and in some lists the word which tied all the other words together was also included. The final memory test asked participants if an item was part of the original critical words or not. Participants were either given a warning about the existence of critical words before looking at the lists or were told about the critical words right before the memory task. They found that pre-warning was more effective than post-warning at preventing false recognition of semantically-related lures. The authors attributed this to the fact that participants who received a pre-warning were primed to identify critical items *while* encoding the word lists and use this information later to reduce false recognition. Yet, they acknowledged that the efficacy of a warning depends on the type of strategies it encourages as well as the types of information in memory that help the participants discriminate between true and false information.

Other research has also found an eyewitness memory benefit for pre-warning as compared to post-warning. In an experiment where participants saw a pick-pocket steel a women's purse, and then read a paragraph of either consistent or misleading information about the crime before a test of their memory, participants who received a pre-warning were found to take longer to read the

post even information and also has higher accuracy scores on critical items (Greene, Flynn, & Loftus, 1982). The researchers suggested that pre-warning led to better memory accuracy due to increased scrutiny of post-event information. While the findings of both Gallo et al. (2001), and Greene et al. (1982) suggest that pre-warning could better protect individuals from misinformation in comparison to post-warning, the effects of pre-warning still remain relatively unstudied compared to effects of post-warning within the context of misinformation paradigms. Furthermore, the effects of pre-warning have yet to be investigated in the context of the repeated retrieval paradigm.

In addition to a lack of research on pre-warning in the context of repeated retrieval and misinformation, there is little understanding about when the mechanisms of warning (pre or post) diminishes the misinformation effect. In a recent review, it was suggested that any effects resulting from post-warnings could only be attributed to the retrieval phase whereas pre-warning could affect the encoding of post-event information, memory retrieval during the final test, or both (Blank & Launay, 2014). For example, pre-warning may modulate encoding-related processes during the post-event information phase by promoting or suppressing attention to misleading post-event information, which may later improve accuracy during the final memory test. In addition, pre-warning may reduce susceptibility to misinformation in a similar manner as post-warning by increasing source monitoring during the final memory test. A final possibility is that pre-warning could modulate both encoding-related processing during the post-event information phase as well as retrieval-related processing during the final memory test.

The present study examines two questions related to misinformation susceptibility. First, what is the effect of warning on eyewitness memory in the context of repeated retrieval and misinformation? Specifically, do pre- and post-warnings similarly reduce the misinformation

effect, or does pre-warning have a greater protective effect and reduce the incorporation of misinformation into memory to a greater extent compared to post-warning? Second, what is the mechanism by which pre-warning and post-warning affect memory performance and how do these types of warnings affect participants' memory strategies? Specifically, does post-warning primarily affect retrieval-related strategies during the final memory test whereas pre-warning affects retrieval-related strategies as well as encoding-related strategies during the post-event information phase?

To address these questions, we conducted a memory experiment designed to mirror the experience of an eyewitness. Mock eyewitnesses first viewed a video of a crime and then later had to remember details about the crime. Between the video and the final memory test, participants were exposed to misinformation (e.g., post-event information) in the form of an audio narrative which recounted the crime but altered some details. Similar to Thomas et al. (2010), we implemented a repeated testing designed where participants also had an initial test of memory before being given post-event information. To investigate the effects of warning on the misinformation effect in this repeated testing context, participants were assigned to one of three warning conditions: no-warning, pre-warning, or post-warning. The sequence of experimental events for the no-warning condition was as follows: Witnessed Event-> Initial Memory Retrieval -> Post-Event Information-> Final Memory Retrieval -> Exit Survey. The sequence of events for the pre-warning condition was as follows: Witnessed Event-> Memory Retrieval -> Warning -> Post-Event Information-> Final Memory Retrieval -> Exit Survey. Lastly, the sequence of events for the post-warning condition was as follows: Witnessed Event-> Initial Memory Retrieval -> Post-Event Information-> Warning -> Final Memory Retrieval -> Exit Survey. Given the findings of Thomas et al. (2010) and Gallo et al. (2001), we predicted that compared to

participants in the no-warning condition, participants in both the pre-warning and post-warning conditions would demonstrate a reduced misinformation effect, as evidenced by an increased accuracy on misleading questions during the final memory test, as well as stable to increased accuracy on control questions. Given past research, we also predicted the pre-warning performance would be better than post warning performance (Gallo et al., 2001; Greene et al., 1982).

To investigate how warning might affect participants' memory strategies, we created an Exit Survey that asked participants about how they interacted with information during the encoding period (post-event audio narrative) and retrieval period (final memory test). The survey consisted of eight questions, each with a Likert scale of 1-7 and a free response section. Regardless of what the question was asking, the lower numbers on the scale were associated with more negative perceptions or lower attention, and the higher numbers on the scale were associated with more positive perceptions or higher attention. We hypothesized that providing participants with a post-warning would modulate strategies during the retrieval period, aligning with the findings of Blank and Launay (2014). In contrast, we predicted that providing participants with a pre-warning would modulate strategies primarily during the encoding period, since early warning about the veracity of post-event information has been shown to diminish misinformation to a larger extent than post-warning in different memory paradigms (Greene et al., 1982).

## Methods

### Participants

Sixty-seven adult participants under the age of 35 were recruited from the Boston area. Eleven participants had to be excluded from analysis for various reasons. Two were excluded because they were older than 35, and five were excluded because there was a technical

malfunction that preventing the completion of the experiment. Lastly, four were excluded because of participant error, such as accidently skipping instructions or falling asleep within the MRI scanner. After exclusions, there were 58 participants, 32 women and 26 men. All participants were right handed, had normal to corrected vision, and no history of brain injury. The mean age was 24 years old (*SD* = 4.42). All participants were provided written informed consent in accordance to the experimental procedures of the Institutional Review Board at Tufts University and they were compensated $20 per hour.

**Materials**

*Witness event (video).* A 22 minute except from the black and white silent film "Rififi" (Vuattoux, & Dassin, 1955) was used as the witness event. The clip portrayed four men committing a burglary in the middle of the night. No participant reported seeing this movie before.

*Initial Memory Test.* Twenty-four questions about specific details from the video were constructed as memory test stimuli for the Initial Memory Test. An example of one of these questions is, "What do the men hang for privacy?". Each multiple-choice question relating to the witness event appeared in the center of the screen, with four answer choices below it. Participants had seven seconds to select their choice by pressing the 1, 2, 3, or 4 keys on the keyboard. Each number corresponded to the order in which the responses where listed, with the response option closest to the top of the screen being marked as 1, and the question closest to the bottom of the screen being marked as 4. After each question relating to the witnessed event, participants were asked to rate their confidence in the answer they just gave. The scale was 1 to

4, with 1 being a complete guess or lowest confidence, to 4 being complete confidence in the answer. Following the confidence question there were a series of arrows that appeared on the screen for eight seconds. The participants had to indicate which direction the arrows were going, either left or right, by pressing the 1 key for left and the 2 key for right. The arrows were included as a baseline period to be used in the fMRI analysis for the Final Memory Test. In addition, arrows were included in the Initial Memory Test to keep the memory tasks consistent. The Initial Memory Test occurred immediately after the witness event.

*Post-Event Information (Audio Narrative).* An audio narrative was used to present post-event information about the witness event after the Initial Memory Test. It was comprised of 130 sentences total, each of which was spoken by a female voice at a standard reading pace and with an even inflection so as to not highlight any one part of the sentence. Of these 130 sentences, there were 24 critical sentences. Specifically, 8 sentences containing consistent information, 8 sentences containing control information, and 8 sentences containing misleading information. Consistent phrases contained details that were accurate regarding the witnessed event (e.g. "The number of men rolling up the rug is **three**"). Control phrases included details from the video in which the critical detail did not support nor contradict what was within the video (*e.g.,* "The number of men rolling up the rug is **a few**"). Misleading phrases included details from the video that had been changed in the narrative (e.g., "The number of men rolling up the rug is **two**"). The critical detail (e.g., two, a few, three) was always present at the end of the sentence. Misleading, neutral, and consistent phrases were counterbalanced across participants. The audio narrative also contained 106 filler sentences which presented information about the video that was not tested in the pre-narrative nor post-narrative cued recall tests. Similar to the initial memory tests,

a series of arrows appeared on the screen in between sentences for 6 seconds and served as a baseline task used for fMRI analyses. Participants had to respond if the arrow was either left facing or right facing.

*Final Memory Test.* This memory test had the same content and structure as the Initial Memory Test described above. Many aspects of the test were exactly the same, such as the question order and the timing of the test, but the order in which the possible responses appeared was counterbalanced across participants and always different from the initial memory test the participant received. In addition, this memory test was conducted within the MRI scanner and participants responded to this test by pressing buttons on a MRI safe button box instead of a computer keyboard. The buttons on the button box were oriented similar to a computer keyboard so that the farthest left button was associated with the first possible response to a question, while the farthest right button was associated with the last possible response listed for a question. All other details of the test were the same.

*Survey.* The Exit Survey was administered to participants at the end of the study (See Appendix A). This survey was created to measure participants' strategies during the encoding of the post-event narrative and during the final memory test after the audio narrative. The survey consisted of eight questions in total, each of which had a 7-Point Likert scale. One was the lower end of the scale indicating low frequency or level of agreement of what the question was inquiring about. Seven was the highest end of the scale indicating high frequency or level of agreement with what the question was asking. Below each scale there was an open-ended section in which participants were able to explain their score or strategy in more detail. The first three of these

questions queried strategies employed during the processing of the audio narrative (encoding). Specifically, they asked “While listening to the audio narrative, how much attention did you pay to the content?”, “While listening to the audio narrative, how frequently did you think back to the video?”, and “While listening to the audio narrative, how frequently did you think back to your answers from the first test”. In addition, there were three questions which queried the strategies employed during the final interview (retrieval). Specifically, they asked “While completing the second test, how frequently did you think back to the video?”, “While completing the second test, how frequently did you think back to the audio narrative?”, and “While completing the second test, how frequently did you think back to your answers from the first test”. Lower scores for these questions were associated with participants ignoring, or not referencing, the corresponding source in their memory while recalling the original crime, while high scores are associated with participants paying greater amounts of attention, or actively referencing, a particular source while searching their memory. In addition to these six questions, we also asked participants to rate the trustworthiness and accuracy of the audio narrative.

**Procedure**

After completing a metal screening and informed consent, participants changed into scrubs and were told to take all metal off their body (hair pins, body jewelry, etc.). They were then given a preliminary metal check to confirm they were free of any metal. After the metal check, participants watched the witness event. They were informed that the video would be 22 minutes long, and that while there would be sounds in the video, there would be no spoken words. Following the video, participants immediately took the cued recalled test on details from the video. Participants were then given a break to use the restroom before getting into the

scanner. Before entering the scan room, participants were checked a second time for metal to make sure they did not pick up any, such as a pen or the locker key, during the first part of the experiment. Each participant had approximately 20 minutes in between the end of the memory test and the start of the audio narrative. Before getting in the scanner participants were fitted with a pair of MRI safe earbuds so they could hear stimuli, and they were given a MRI safe button box in to make their responses. Once in the scanner participants would be instructed to keep their eyes open and look at a fixation cross in the middle of the screen. The first task within the scanner was audio calibration so the participant would be able to hear stimuli while the scanner was running. Following the calibration, participants in the **pre-warning condition** were cautioned about the validity of the audio narrative they were about to hear. Specifically, participants heard, "You will have to answer questions regarding the video you previously watched for a second time. We will play a narrative of that video; however, we are uncertain as to the source of the narrative. Therefore, we were unable to verify the accuracy of the narrative. As such, base your answers only on what you saw in the video, and not on what you hear in the narrative.". The participants in the **post- warning condition** and the **no-warning condition** went directly from the audio calibration to listening to the audio narrative. After the audio narrative, participants in the **post-warning condition** were cautioned about the validity of the audio narrative they just heard, before moving on to the final memory test. Specifically, participants heard, "You will have to answer questions regarding the video you previously watched for a second time. We just played a narrative of that video; however, we are uncertain as to the source of the narrative. Therefore, we were unable to verify the accuracy of the narrative. As such, base your answer only on what you saw in the video, and not on what you heard in the narrative.". The **pre-warning condition** and the **no-warning condition** received no warnings following the

audio narrative, before they took the final memory test. Once these scans were finished, participants were removed from the scanner and given the Exit Survey. They were asked to elaborate on their answers for the first Exit Survey, but it was not required. Once the survey was completed participants were debriefed and compensated for their time.

**Analysis**

Past research has found that participants who receive a warning about misinformation *after* encountering misinformation show a lower RES effect compared to participants who received no warning (Thomas et. al., 2010). For clarification, since all conditions within this experiment involved an initial test of memory, and RES effect is the difference between subject receiving initial testing or not, we did not look at the RES effect but the misinformation effect. Still, even within past misinformation research involving only testing after post-event information, post-warning has been found to significantly diminish the misinformation effect (Blank & Launay, 2014). In order to look at the misinformation effect within the present study, we used an analysis method outlined in Blank and Launay's (2014) meta-analysis. Traditionally, the misinformation effect is defined as comparing the difference of the control accuracy and misleading accuracy across conditions. Yet, Blank and Launay (2014) suggest that this definition does not account for the possibility that warning could make participants more conservative in reporting, thus effecting performance across trial types. For example, when looking at the difference between control accuracy and misleading accuracy, a decrease in the difference could come from higher misleading accuracy *or* lowered control accuracy. This means that we could attribute a diminished difference between control accuracy and misleading accuracy as an improved misleading accuracy when in reality the misleading accuracy stayed the same and the

control accuracy worsened. This would mean that participants did not get any better at misleading questions, but instead got worse at control questions. In order to avoid this problem, Blank and Launay (2014) proposed examining control accuracy and misleading accuracy separately so that how the warning is affecting both question types are clear.

Additionally, fMRI data and confidence data collection were elaborated upon within this study's methods section in order to best recount the procedure of this experiment. Of note, neither will be discussed any further.

## Results

### Effect of Warning on Eyewitness Memory

In the present study we implemented Blank and Launay's (2014) methodology by conducting one-way ANOVAs on misinformation accuracy and control accuracy. First, a one-way between subjects ANOVA was conducted to compare the effects of warning on control question accuracy in the no-warning, pre-warning, and post-warning condition. There was a significant effect of warning on control question accuracy for the three conditions ($F(2,55) = 3.75$, $p = .03$) (Figure 1). Pair-wise comparisons of each variation of condition on control accuracy were conducted. Participants within the no-warning condition were less accurate on control questions ($M = .57$, $SD = .25$) compared to the accuracy of participants within the post-warning condition ($M = .77$, $SD = .20$), $t(38) = -2.84$, $p = .01$. There was not a significant difference between accuracy of participants in the no-warning condition and accuracy of participants in the pre-warning condition on control questions, $t(34) = -.66$, $p = .52$, nor between the accuracy of participants in the pre-warning condition compared to the accuracy of participants in the post-warning condition, $t(38) = -1.90$, $p = .07$ .

A one-way between subjects ANOVA was conducted to compare the effects of warning on misleading question accuracy in the no-warning, pre-warning, and post-warning condition. There was a significant effect of warning on misleading question accuracy for the three conditions ($F(2,55) = 5.36$, $p = .01$) (Figure 2). Pair-wise comparisons of each variation of condition on control accuracy were conducted. Participants within the no-warning condition were less accurate on misleading questions ($M = .36$, $SD = .21$) compared to participants in the pre-warning condition ($M = .59$, $SD = .26$), $t(34) = -2.99$, $p = .001$. In addition, participants within the no-warning condition we less accurate on misleading questions ($M = .36$, $SD = .21$) compared to participants in the post-warning condition ($M = .57$, $SD = .25$), $t(38) = -2.91$, $p = .01$. There was no significant difference between the accuracy of participants in the pre-warning condition and post-warning condition on misleading questions, $t(38) = .25$, $p = .81$.

In addition to comparing accuracy of control questions and accuracy of misleading questions separately, we looked at misinformation selection (Figure 3). Misinformation selection is the proportion of misleading responses given for questions connected to misleading details within the post-event narrative. This analysis of data means allowed us to test the degree to which participants are reporting the misinformation that is being told to them. To compare misinformation selection in the no-warning, pre-warning, and post-warning conditions we conducted a one-way between subjects ANOVA. There was a significant effect of warning on misinformation selection for the three conditions ($F(2,55) = 6.42$, $p = .003$). Pair-wise comparisons of each variation of condition on misinformation selection were conducted. Participants within the no-warning condition were more likely to choose the misleading answer on misleading questions ($M = .51$, $SD = .21$) compared to participants within the pre-warning condition ($M = .29$, $SD = .20$), $t(34) = 3.28$, $p = .002$. Similarly, participants within the no-warning

condition were also more likely to choose the misleading answer on misleading questions ($M$ = .51, $SD$ = .21) compared to participants within the post-warning condition ($M$ =.30, $SD$ = .23), $t(38) = 3.05$, $p$ =.004. There was no significant difference between the participants in the pre-warning and post-warning conditions in regard to misinformation selection, $t(38) = -.20$, $p$ = .85. These findings suggest that, similar to pre-warning and post-warnings effects on misleading question accuracy, that pre-warning and post-warning modulate misinformation to a similar extent.

**Effect of Warning on Strategy**

A total of four possible strategies involving source monitoring were identified. Two connecting to encoding strategy (question 1 and 2) and two connecting to retrieval strategy (question 6 and 7). Each of these strategies are outlined below and are followed by the relevant findings.

**Encoding Strategy**

The first strategy related the encoding period (audio narrative) was operationalized as higher scores on both question 1 (“While listening to the audio narrative in the scanner, how closely did you pay attention to what you heard?”) and question 2 (“While listening to the audio narrative in the scanner, how frequently did you think back to the video?”). This means that participants were reporting thinking about the audio narrative and thinking back to the video during the audio narrative very frequently. This would be considered enhancement of source monitoring as the participants would be seen as processing these two sources of information and possibly even comparing and contrasting the information in real time. The second strategy during the encoding period was operationalized as a low score on question 1. This would be considered

suppression of the audio narrative because the participant would be reporting they ignored what they were hearing. Thus, their strategy would be to not encode, or simply *suppress,* the information once they knew it could be incorrect information.

We ran one-way ANOVAs for question 1 and question 2 with the function of the analysis being condition (no-warning, pre-warning, post-warning). In the one-way ANOVA looking at the effect of condition on question 1 scored we found no significant differences, $F(2,55) = .09$, $p = .92$ (Figure 4). Similarly, we found no significant difference when looking at the effect of warning in question 2 scores, $F(2,55) = .39$, $p = .68$ (Figure 5). This suggests that regardless of condition, participants responded interacting with the data during the encoding period in similar ways.

Supplementary to these findings, the majority of participants across conditions who filled out the free response for question 1 and question 2 reported comparing what they heard in the audio narrative to what they saw in the video (Table 1). For example, one participant wrote "I listened and thought about the movie" in response to question 1. Importantly, participants commented on their strategy in either the free response for question 1 or question 2, showing that no one strategy was associated with either question 1 or question 2 but instead across both questions.

**Retrieval Strategy**

The first strategy related to the retrieval period (final memory test) was operationalized as a high score on question 6 and 7. Similar to the first strategy related to the encoding period, this would be seen as enhancement as the participant would be reporting increased source monitoring and would possibly be comparing the information from the audio narrative and the video. The second strategy related to the retrieval period was operationalized as a high score on question 7

("While completing the multiple-choice test in the scanner, how often did you think back to the audio narrative?") and a low score on question 6 ("While completing the multiple-choice test in the scanner, how often did you think back to the video?"). While this score would typically be considered enhanced source monitoring because the participant is reporting thinking frequently back to the video while ignoring information from the post-event information, the fact that they do suppress the audio narrative means that it can be considered suppression since the participant would be seen as trying to block out the narrative information.

A one-way between subjects ANOVA was conducted to compare the effects of warning on question 6 scores in the no-warning, pre-warning, and post-warning condition. There was a significant effect of warning on question 6 scores for the three conditions ($F(2,55) = 4.53$, $p = .02$) (Figure 6). Pair-wise comparisons of each variation of condition were then conducted for question 6. Participants within the no-warning condition reported lower scores ($M = 5.28$, $SD = 1.74$) than the participants within the pre-warning condition ($M = 6.50$, $SD = .71$), $t(34) = -2.76$, $p = .01$. In addition, participants within the no-warning condition reported lower scores ($M = 5.28$, $SD = 1.74$) than participants in the post-warning condition ($M = 6.13$, $SD = 1.13$), yet the difference was not statistically significant $t(38) = -1.88$, $p = .07$. While there was no significant difference between the no-warning and post-warning conditions, the fact the difference was approaching significance indicants that participants in the post-warning condition were reporting thinking back to the video at higher frequencies than those who did not receive a warning. There was not a significant difference between participants' scores in the pre-warning condition and the post-warning condition, $t(38) = 1.19$, $p = .24$.

Like with question 6, a one-way between subjects ANOVA was conducted to compare the effect of warning on question 7 scores in the no-warning, pre-warning, and post-warning

condition. There was a significant effect of warning on question 7 scores for the three conditions ($F(2,55) = 4.14$, $p = .02$) (Figure 7). Pair-wise comparisons of each condition for question 7 were conducted. Participants in the no-warning condition scored higher on question 7 ($M = 5.72$, $SD = 1.67$) than the participants in the pre-warning condition ($M = 4.11$, $SD = 1.99$), $t(34) = 2.62$, $p = .03$. In addition, participants in the no-warning condition also scored higher on question 7($M = 5.72$, $SD = 1.67$) than participants in the post-warning condition ($M = 4.64$, $SD = 1.49$), $t(38) = 2.16$, $p = .04$. There was no significant difference between participants in the pre-warning condition and participants in the post-warning condition's scores on question 7, $t(38) = -.95$, $p = .35$.

In addition to the score responses, some participants also gave free responses to question 6 and 7. These free responses seem to be consistent with findings above that demonstrate participants within the warning conditions source monitor to greater degrees than participants who did not receive a warning. For question 6, less than half of participants who wrote a free response responded that they thought back to the video during the final test of memory (Table 1). Yet, the majority of participants in the pre-warning condition and the post-warning condition commented about thinking back to the video when responding to final memory test questions (Table 1). For example, one participant within the post-warning condition stated, "I used the video as my first source for answers". Furthermore, free responses for question 7 also map on well to the memory accuracy results for question 7. More than half of participants in the no-warning condition report using the audio as their *main form* of memory for responding to final memory test questions, while no participants from their warning condition report doing so.

## Discussion

The present study investigated two aspects of eyewitness testimony that have been overlooked in past research. The first was the effect of pre-warning on the misinformation effect in the repeated testing paradigm as compared to post-warning and no warning. As expected from past literature, participants who received a pre-warning had a diminished misinformation effect compared to participants who received no warning. Yet, the degree in which the misinformation effect was lowered was not statistically different from the degree to which the misinformation effect was lowered by post-warning. This was surprising given past research suggesting that pre-warning had stronger effects of preventing effects of misinformation than post-warning (Gallo et al., 2001; Greene et al., 1982). The second novel aspect of the current study was its investigation of how warning modulates participants' interaction with post-event information. Interestingly, all participants, regardless of which condition they were in, reported similar levels of attending to the audio narrative and thinking back to the witness event (the video) during the encoding phase (audio narrative). In other words, they all reported source monitoring. Furthermore, participants who received any type of warning reported statistically higher levels of source monitoring during the retrieval period. Specifically, warnings were associated with paying more attention to the witness event (video) and less attention to the post-event information (audio-narrative) during retrieval. While the reports of source monitoring during both periods of the experiment are not surprising given the speculation of past literature (e.g., Johnson et al., 1993), it is significant to see such clear subjective reports of source monitoring from the participants themselves. Together, these results provide novel findings that both pre-warning and post-warning modulate the misinformation effect in a similar manner. Additionally, participants' subjective reports suggest that both types of warning appear to modulate memory strategies during the final

memory test (retrieval period), and that participants invoke a strategy of source monitoring to reduce the misinformation effect.

**The Misinformation Effect and Warning**

In this study, the misinformation effect was investigated by looking at accuracy of misleading questions and accuracy of control questions separately in order to compare the possible benefits of warning (increase in misleading question accuracy) to the possible costs of warning (decrease in control question accuracy). It is important to note that we found no negative effects on control question accuracy for either warning type. In fact, pre-warning increased control accuracy by approximately 10% compared to no-warning control accuracy and post-warning increased control accuracy by approximately 23% compared to no-warning control accuracy. While the 10% increase for pre-warning was not significant, this novel trend is the first to suggest that pre-warning could benefit control accuracy in an eyewitness study. The degree to which post-warning control accuracy was increased in the current study is greater than what has been observed in past research (Thomas et al., 2010; Wulff Master's Thesis, 2019), raising questions about the factors that might influence the effect of post-warning on control accuracy. In addition, this is a novel finding for pre-warning, as past literature only looked at the difference between consistent post-event information and misleading post-event information and had only ever found warning improved accuracy of misleading questions (Greene et al., 1982). Given that this is the first time these findings have been observed for either warning, future research will need to be conducted to understand these relationships before anything can be extrapolated from these findings.

We also found that pre-warning and post-warning increased misleading question accuracy compared to misleading question accuracy of participants in the no-warning condition (~23% and ~21%, respectively). Our finding that pre-warning increased accuracy of misleading questions is consistent with prior experiments which have implicated that pre-warning lowers misinformation selection in other paradigms (Greene et al., 1982). Yet, this finding is also novel because it is the first time that pre-warning has been shown to lower the misinformation effect within an eyewitness experiment, more specifically one with a repeated testing paradigm. For post-warning, our findings align with prior research that found that post-warnings can lower the misinformation effect in the context of repeated testing paradigms (Thomas et al., 2010). Still, what is particularly interesting when comparing the effect of pre-warning and post-warning on the misinformation effect in the current study is that both types of warning increased participants' accuracy of misleading questions by comparable amounts. This finding suggests that warning, no matter when it is given, will similarly improve the misinformation effect in regard to misleading question accuracy.

In addition to participants' accuracy on control and misleading questions, we also looked at misinformation selection for misinformation questions. Similar to prior research, we found that both pre-warning and post-warning diminished the rate at which participants produced misinformation significantly (Thomas et al., 2010). These findings align with past findings, in addition to providing novel evidence that pre-warning has a beneficial effect on misinformation selection in an eyewitness memory study. Additionally, these findings mirror the misleading accuracy findings within the current study and support the idea that pre-warning and post-warning are interacting with misleading information in similar ways. Taken together, the findings of this research suggest that with repeated testing, pre-warning and post-warning not

only improve overall accuracy of misleading questions, but also diminish the amount of misleading information produced in the final recall test.

**Encoding Strategy**

Our finding that there is no difference with respect to participants' self-reported strategies between conditions during the encoding period contradicts past research which attributed pre-warning's effect on misinformation to participants paying more attention to the information they received after the warning (Gallo et al., 2001; Greene et al., 1982). There was no significant difference between conditions for question 1 or question 2 of the Exit Survey. That means that participants were reporting paying attention to the audio narrative and thinking back to the video during the audio narrative at similar levels across conditions. Current findings suggest that participants across all conditions report interacting with post-event information in a similar manner.

In addition to the memory accuracy, the subjective free responses helped to illuminate how participants were thinking about their actions. Across the board for encoding-related free responses, participants reported cross-comparing answers and thus acknowledged the two different sources from which information could be coming from. In addition, participants reported this cross-comparison strategy at similar frequencies, a response which mirrors the score findings of question 1 and 2 which showed there was no difference across condition in how participants rated how much they interacted with the audio narrative and thought back to the video during post-event information encoding. While we do not know if these self-assessments accurately reflect what participants are actually doing, it is still promising to see similar results across different subjective data. Additionally, these explicit comments on thinking back and forth

between two sources supports the idea that participants are source monitoring during the encoding period.

Taken together, these findings help narrow down our understanding of how and when warnings modulate participants' memories. Past research hypothesized that pre-warnings have the ability to modulate encoding, retrieval, or both (Blank & Launay, 2014). The current findings that there was no difference in how participants across conditions reported interacting with data during the encoding phase suggests that pre-warning is not changing how participants interact with post-event information as they hear it. This suggests that pre-warnings likely modulate memory during the final memory test (retrieval period).

**Retrieval Strategy**

While there was no significant difference between conditions for encoding strategy, we did find significant differences across conditions for retrieval strategy. For question 6, which specifically quarries if participant were thinking about the witnessed event (video) during the final memory test, both pre-warning and post-warning participants reported thinking about the video at higher rates than participants within the no-warning condition. While the difference between no-warning and pre-warning were statistically significant, the difference between no-warning and post-warning only approached significance. Despite this lack of statistical significance across waring types, these findings still support the idea that both warning types led participants to source monitor by specifically increasing how frequent they thought back to the video source while answering questions on the final memory test. This same trend of warning increasing source monitoring was also shown in the findings relating to question 7. Question 7 asked participants how frequently they were thinking back to the audio narrative during the final

memory test. Participants within the pre-warning and post-warning conditions reported thinking back to the audio narrative at significantly lower rates than participants within the no-warning condition. While participants in the warning conditions reported suppressing their memory of the audio narrative, this response can be considered source monitoring because when taken together with the fact they reported increasing how frequently they think about the video, it is clear that participants reported actively distinguishing between two sources of memory.

These findings suggesting the use of source monitoring are also supported by the free response section of question 6 and question 7. Compared to participants who did not receive a warning, participants in the warning conditions reported thinking of the video more and using the audio narrative to a smaller degree when making final memory test responses. For example, one participant from the pre-warning condition stated "I used the video as my first source for answers" for question 6, and "I used the audio narrative to supplement the questions…" for question 7. While this does not mean that participants within the warning conditions were completely suppressing the memory of the audio narrative, this suggests that no participants within the warning condition were treating the audio narrative as the only source of information from which to retrieve memory. In addition, these responses demonstrate that participants in the warning conditions were able to distinguish these two events as separate memories and attempt to draw details from only one specific source.

**Overall Interpretations**

The current study represents an important step in exploring the relationships between warning, the misinformation effect, and memory strategies. This study is the first to examine the effects of pre-warning and post-warning in a repeated testing paradigm. As such, we were able to

confirm past findings that post-warning diminished the misinformation effect and misinformation selection (Thomas et al. 2010). We were also able to provide novel evidence that pre-warning lowers the misinformation effect and misinformation selection in the repeated testing paradigm. Importantly, our finding that both of these warning types lower the misinformation effect and misinformation selection by similar degrees is the first to suggest that warnings reduce misinformation effect in the context of repeated testing regardless of when they are provided. Furthermore, this is also the first study which uses subjective measures to inquire about how and when participants interact with information during the experiment in order to gain insight into how warnings modulate participant behavior. Our findings that regardless of condition participants report similar actions during the encoding period is inconsistent with past research suggesting that pre-warning increases attention or could modulate attention during the encoding period (Gallo et al., 2001; Greene et al., 1982; Blank & Launay, 2014). Yet, this difference between past research and the current study could be possibly accounted for since this study has a significantly different methodology compared to past research. In addition to providing insight into perceived encoding strategy, this study also provides the first evidence that both pre-warning and post-warning participants report source monitoring during the retrieval period. These findings, while consistent with past research that has discussed the possible effects of post-warnings on retrieval, provide important evidence that the critical period of behavior modulation is during the retrieval period since both pre-warning and post-warning conditions reported similarly and significantly changing their behavior during this time in the experiment (Blank & Launay, 2014).

Together, these results suggest that warnings are a powerful way to influence both eyewitness memory strategies and eyewitness memory accuracy. Specifically, warning

eyewitnesses about the veracity of post-event information either before or after exposure to that information can enhance source monitoring at retrieval and improve memory accuracy. These findings have several practical implications for eyewitness interrogations. First, they suggest that warning is a viable option for lowering the detrimental misinformation effect. While warning does not eliminate the misinformation effect, the current study shows that both pre-warning and post-warning increase misleading question accuracy at no cost to control accuracy. Additionally, both pre-warning and post-warning lower the amount of misinformation that is produced in response to misleading questions. Together, using warning within eyewitness interrogations could mean that investigators would get testimony that is not only more accurate, but also contains less misinformation.

**Limitations and Future Directions**

While this study is in line with earlier research and provides exciting novel findings in regard to warnings, it has limitations that should be addressed in future research. First and foremost, this study only examined memory strategy via subjective reports. This means that participants recounted what they believed happened, while what actually happened could be something completely different and outside of their awareness. Being able to understand one's cognitive processes is called metacognition, and past confidence research has demonstrated that people's ability to identify internal memory processes varies (Perfect & Hollins, 1996). Thus, while the current study's findings present interesting patters in regard to strategy, poor metacognition might be preventing participants from accurately assessing the true processes they are using. For example, our subjective data suggests that across conditions there was no difference in how participants interacted with information at encoding, yet objective measures could find another pattern. Thus, an important area of future research is to use objective

measures in tandem with subjective measures to see if patterns are consistent across data types. Neuroimaging, such as fMRI, could be used as an objective measure since it can look at activity within specific regions of the brain and provide insight into what cognitive processes participants are using. Notably, the current study is a part of a larger study where part of the current data was collected within an fMRI machine, thus neural data of this type is available and must simply be analyzed in the future. Another objective measure that could be investigated are response times (RTs). Earlier research has shown that RTs can be indicators of attention and have been used to explain retrieval fluency of details related to more recently acquired memories compared to older memories (Thomas et al., 2010). This piece of research only looked at the effects of an initial test of memory with post warning, thus it would be beneficial to investigate if other warnings change the time participants spend on answering questions on the final test of memory, or how much time they spend making confidence decisions.

Another limitation of this study that could be explored with future research is to have younger adults and older adults complete the study. The current study specifically recruited participants under the age of 35, but Bulevich and Thomas (2012) have found that the amount of support and the type of memory test given to a witness effects eyewitness accuracy depending on the mock witness' age. This means that current findings within our data are not likely applicable to older age groups. Future studies should investigate the effects of warnings on eyewitness memory in older adults, as there is no uniform age of eyewitnesses, and in order to best protect witnesses' memory from misinformation we need to know if we should have different warning protocols based on the age of the eyewitness.

Lastly, another future direction for this research would be to investigate pre-warning and post-warning within different paradigms. Past research has shown that initial testing has different

effects on final memory test accuracy based on how post-event information is presented (Chan & LaPaglia, 2013). Since post-event information is not presented to eyewitnesses in one standard way, this suggests that we need to look into different ways of presenting post-event information within lab settings and how different types of warnings effect these different paradigms. This paradigm consciousness would allow researchers to best understand how warning would affect all the various situations that an eyewitness could encounter.

**Conclusion**

In conclusion, the current study provides initial evidence that pre-warning and post warning diminish the misinformation effect, without compromising control accuracy, by similar degrees within a repeated testing paradigm. Additionally, subjective reports from participants suggest that both warning types influence source monitoring to a statistically significant degree during the retrieval phase. Future research should aim to confirm this relationship with objective measures, such as neuroimaging and response times, as well as investigate any age differences that might exist. Despite the limitations of the current study, it represents a novel step in understanding the effects of pre-warning and post-warning and how they can be used to protect eyewitness memory accuracy.

References

Blank, H., & Launary, C. (2014). How to protect eyewitness memory against the misinformation effect: A meta-analysis of post-warning studies. *Journal of Applied Research in Memory and Cognition, 3,* 77-88.

Bulevich, J.B., & Thomas, A. K. (2012). Retrieval effort improves memory and metamemory in the face of misinformation. *Journal of Memory & Language, 67,* 45-58.

Chan, J. C. K., & LaPaglia, J. A. (2013). Impairing existing declarative memory in humans by disrupting reconsolidation. *Proceedings of the National Academy of Sciecnes of the United States of America, 110,* 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.

Frenda, S. J., Nicholas, R. M., & Loftus, E. F. (2011). Current issues and advance in misinformation research. *Current Directions in Psychological Science, 20(1),* 20 -23.

Gallo, D., Roediger, H., & McDermott, K. (2001). Associative false recognition occurs without strategic criterion shifts. *Psychonomic Bulletin & Review, 8,* 579-586.

Gordon, L. T., & Thomas, A. K. (2014). Testing potentiates learning in the misinformation paradigm. *Memory and Cognition, 42,* 186-197.

Greene, E., Flynn, M. S., & Loftus, E. F. (1982). Inducing resistance to misleading information. *Journal of Verbal Learning and Verbal Behavior, 21,* 207-219.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. (1993) Source Monitoring, *Psychology Bulletin, 114(1),* 3-28.

Innocence Project (2016). Retrieved April 5, 2019, (http://www.innocenceproject.org/).

Lindsay, D. S., Allen, B. P., Chan, J. C. K., & Dahl, L. C. (2004). Eyewitness suggestibility and

source similarity: Intrusions of details from one event into memory reports of another event. *Journal of Memory and Language, 50,* 96-111.

Loftus, E. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning and Memory, 12,* 361-366.

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,* 19-31.

Perfect, T. J., & Hollins, T. S. (1996). Predictive feeling of knowing judgements and postdictive confidence judgements in eyewitness memory and general knowledge. *Applied Cognitive Psychology, 10,* 371-382.

Thomas, A. K., Bulevich, J. B., & Chan, J.C.K. (2010). Testing promotes eyewitness accuracy with a warning: Implications for retrieval enhanced suggestibility. *Journal of Memory and Language, 63,* 149-157.

Vuattoux, R. G. & Dassin, J. (1955). *Rififi*. France: Pathé-Consortium Cinéma.

Wulff, A., (Master's Thesis Under Review). Warning placement affects eyewitness accuracy and confidence in the presence of misinformation.

Appendix A

## Exit Survey

**Part 1.** For the following questions, please think about the the audio narrative that you heard while in the scanner. For each question, feel free to explain your answer in the space below labeled "Explain".

**(1) While listening to the audio narrative in the scanner, how closely did you pay attention to what you heard?**

1 2 3 4 5 6 7
Paid No Attention Paid Close Attention

Explain:

**(2) While listening to the audio narrative in the scanner, how frequently did you think back to the video?**

1 2 3 4 5 6 7
Never Very Frequently

Explain:

**(3) While listening to the audio narrative in the scanner, how frequently did you think back to your answers from the first test?**

1 2 3 4 5 6 7
Never Very Frequently

Explain:

**(4) Rate the overall *accuracy* of the audio narrative that you heard in the scanner.**

1 2 3 4 5 6 7
Not Accurate Very Accurate

Explain:

**(5) Rate the overall *truthfulness* of the audio narrative that you heard in the scanner.**

1 2 3 4 5 6 7
Not Truthful Very Truthful

Explain:

**Part 2.** For the following questions, please think about the test that you completed while in the scanner. For each question, feel free to explain your answer in the space below labeled "Explain".

**(1) While completing the multiple-choice test in the scanner, how frequently did you think back to the video?**

1 2 3 4 5 6 7
Never Very Frequently

Explain:

**(2) While completing the multiple-choice test in the scanner, how frequently did you think back to the audio narrative?**

1 2 3 4 5 6 7
Never Very Frequently

Explain:

**(3) While completing the multiple-choice test in the scanner, how frequently did you think back to the answers that you provided on the first test?**

1 2 3 4 5 6 7
Never Very Frequently

Explain:

**Table 1.** Free Responses in relation to strategies used while interacting with post-event information and during final memory retrieval. S.M. = Source Monitoring, Q1 = Exit Survey Question 1, Q2 = Exit Survey Question 2, Q6 = Survey Question 6, Q7 = Survey Question 7.

**Source Monitoring Free Response**

| | Encoding | Retrieval | |
|---|---|---|---|
| Warning Condition | Proportion reporting S.M. (Q1/Q2) | Proportion reporting using video (Q6) | Proportion reporting using audio as primary source (Q7) |
| No Warning | .75 | .36 | .64 |
| Pre-Warning | .83 | 1 | 0 |
| Post-Warning | .92 | .82 | 0 |

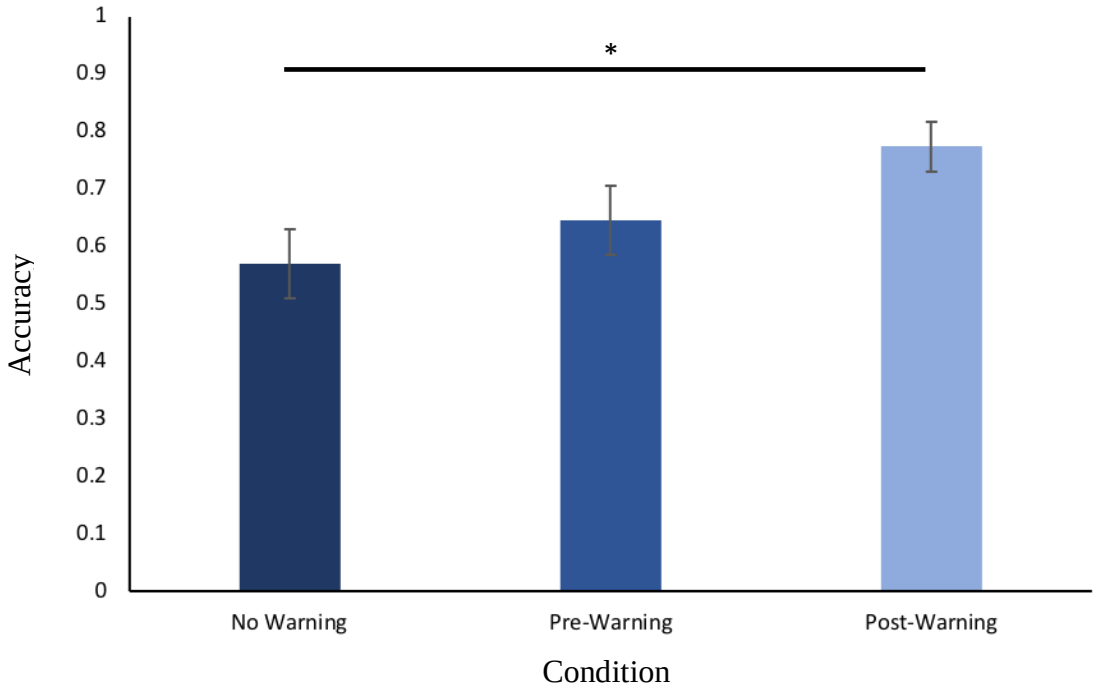


**Figure 1.** Relationship between Control Question Accuracy and Condition Type

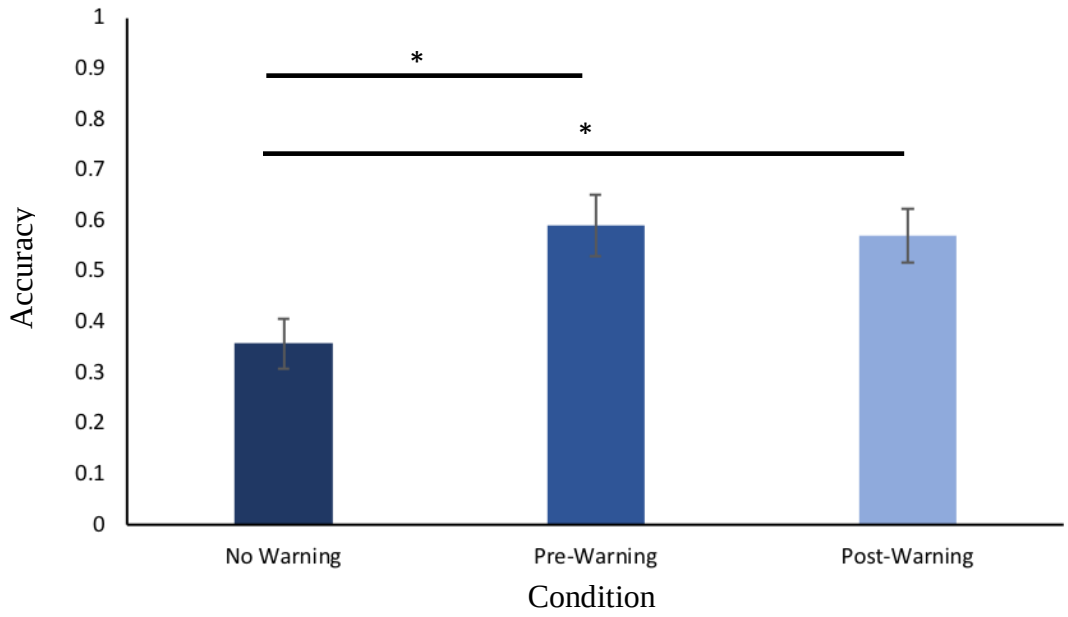


**Figure 2.** Relationship between Misleading Question Accuracy and Condition Type

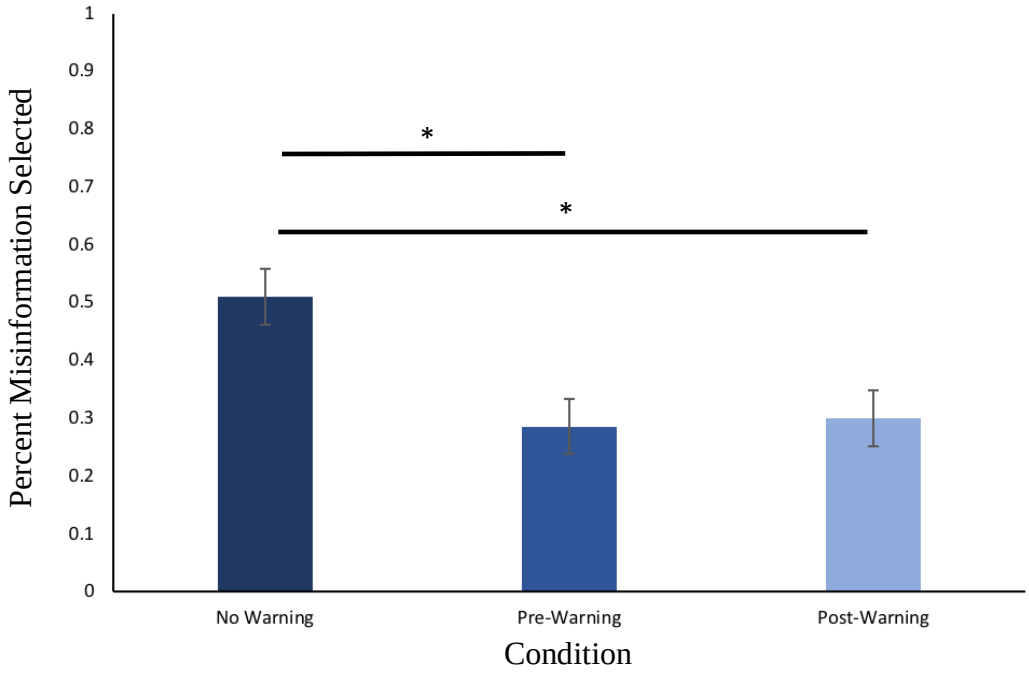


**Figure 3.** Relationship between amount of Misinformation selection on misleading questions and Condition Type

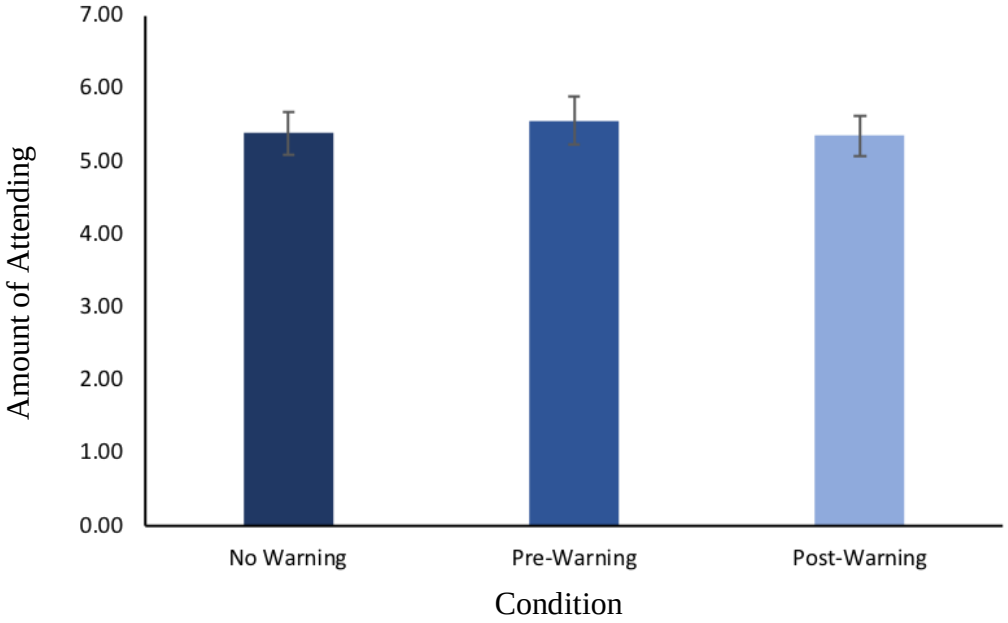


**Figure 4.** Question 1: Relationship between amount of attention paid to the audio narrative during the audio narrative and condition type

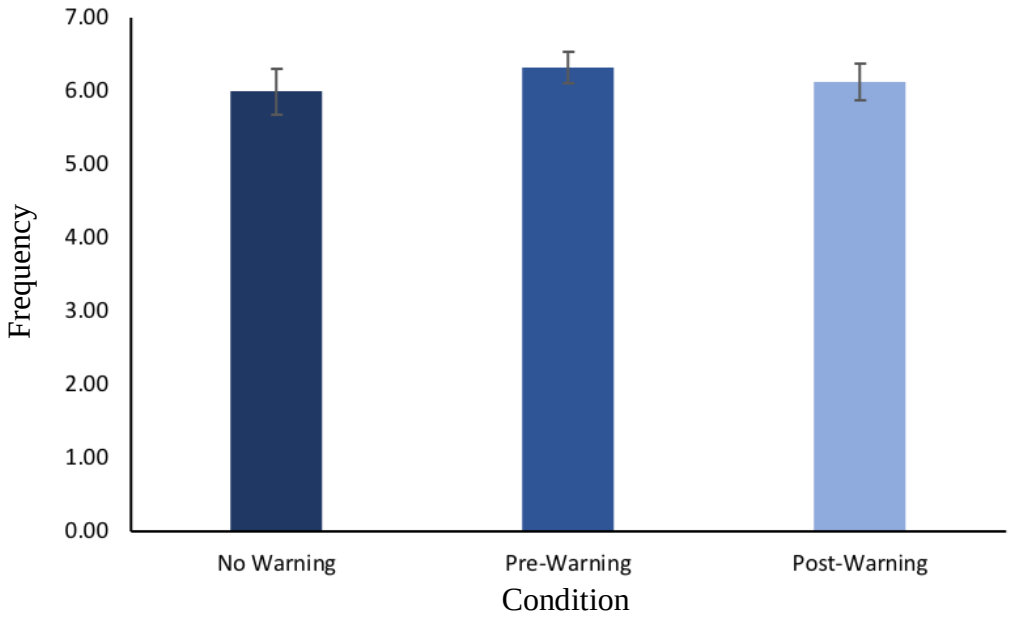


**Figure 5.** Question 2: Relationship between frequency of thinking back to the video during the audio narrative and condition type

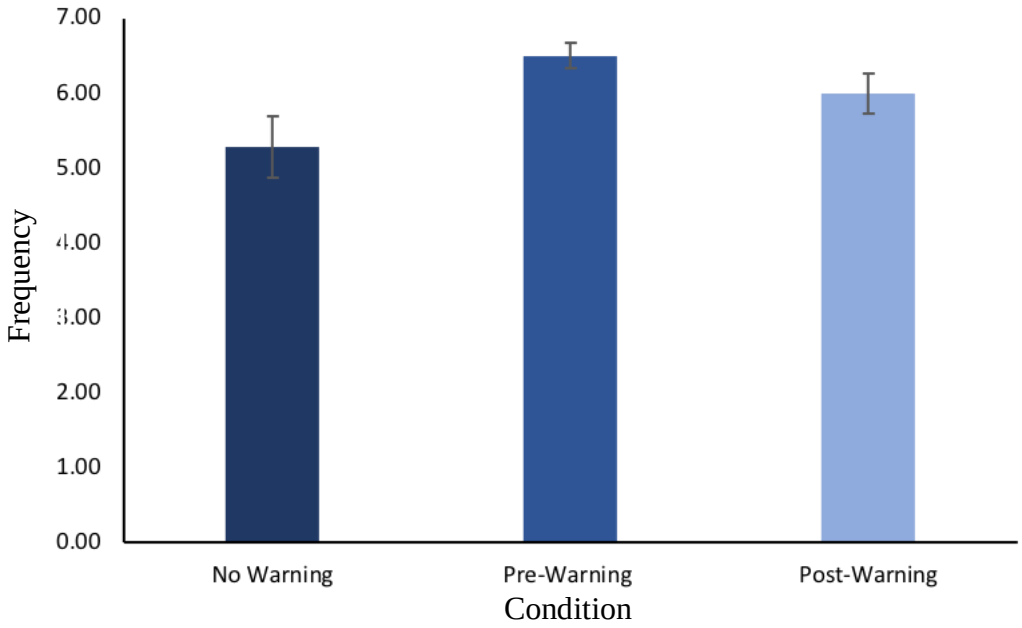


**Figure 6.** Question 6: Relationship between frequency of thinking back to the video during the second memory test and condition type

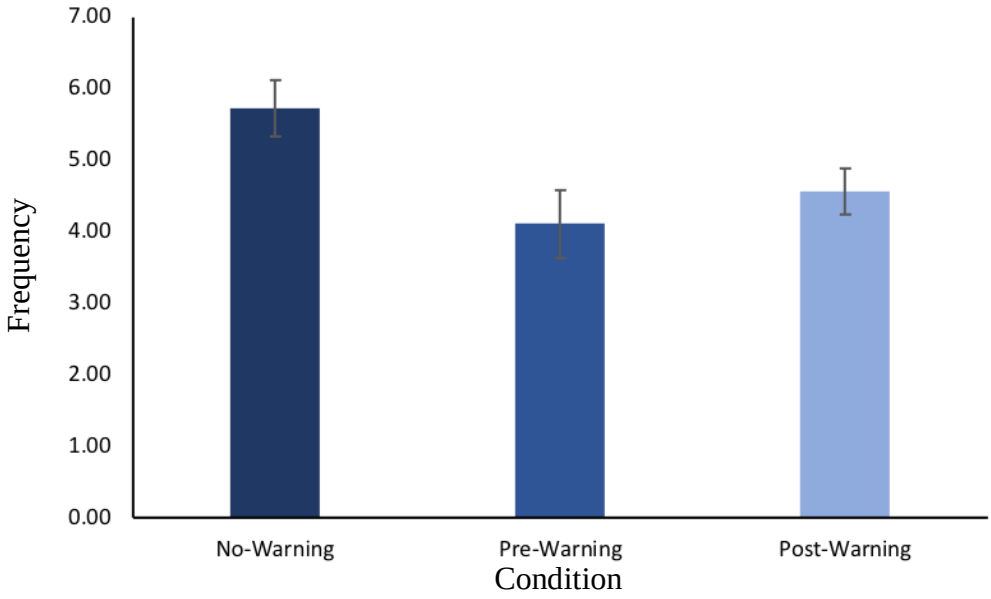


**Figure 7.** Question 7: Relationship between frequency of thinking back to the audio narrative during the second memory test and condition type