Heed the Warning! How Reducing Eyewitness Misinformation with Postwarnings May be Extended to the Cognitive Interview

Chandler Tyler

Advisor: Chad Dodson
Second Reader: Sophie Trawalter

University of Virginia
Abstract

One problem with eyewitnesses is that their memory is not always accurate. False details, or misinformation, may be unintentionally incorporated into an eyewitness’s account of a crime, thereby reducing the validity of their testimony. The present study seeks to mitigate this misinformation effect by explicitly warning participants. However, instead of testing their memory with conventional laboratory memory tests, we used the more ecologically valid cognitive interview procedure. Participants were 48 college students who watched a video of a robbery, read a narrative that suggested false details about the crime, and then experienced the cognitive interview to evaluate their memory. Results indicated that eyewitnesses in the warning condition reported significantly less misleading information during the cognitive interview than those who were not warned.
False memories occur when people try to recollect past experiences but unintentionally misremember the facts. This cognitive phenomenon can be deleterious in judicial settings where eyewitnesses are relied upon to recall what happened during a crime. How accurately can eyewitnesses actually remember what they saw or what they think they saw? More importantly, to what extent can they resist false information stemming from other sources? This study investigated whether warnings can reduce reports of misinformation when memory is interrogated with a relatively ecologically valid procedure.

Eyewitnesses often experience a bombardment of extra information after witnessing a criminal act (Echterhoff & Hirst, 2005). Additional details may come from other witnesses or news media that may over-embellish or unknowingly alter the story as a result of interacting with eyewitnesses (Echterhoff, Groll, & Hirst, 2007). The concern is that these third party sources may inadvertently describe things that are, in fact, not true at all; thereby indirectly and detrimentally influencing the memory of an eyewitness. For instance, if the original eyewitness overhears another person talking about how the thief was wearing black gloves, she might assume this to be true even though she did not actually observe this detail. More importantly, this incorporation of false information now becomes the sworn testimony of the witness, and this is technically not considered perjury because she does not realize her testimony is false (Wise & Safer, 2004). The recent implementation of DNA testing in the criminal justice system has proven that thousands of innocent people have been convicted. Considering that eyewitness error accounts for at least half of all wrongful convictions in the United States, it is critical to
understand how memory may be altered and distorted by these sources of misinformation (Wise et al., 2004).

Post-event misinformation describes the kind of specious information that is presented after an event (e.g. crime) occurs (Echterhoff & Hirst, 2005). Although misinformation is a realistic problem that psychologists and legal experts are well aware of, we still know little about how to reduce the harmful effects of misleading information. One strategy involves warning eyewitnesses about any potentially false information that they have encoded. These warnings can be given before or after the misinformation is received, but it is obviously difficult to submit a prewarning to a witness before they encounter misinformation from co-witnesses or the like (Echterhoff et al., 2005). Therefore, the study of postwarnings proves to be more effective and practical, especially given the time delay between witnessing the crime and reporting to or interviewing with law officers.

There are a variety of ways in which someone can be warned of impending false information. Social, or implicit, postwarnings usually identify third-party sources as having low-credibility, like an intoxicated cowitness or someone with a conflict of interest (Echterhoff et al., 2005). In comparison, explicit warnings directly describe the additional information as being potentially false and ask that participants “monitor” their recollection of that source (Echterhoff et al., 2005). Explicit postwarnings may be advantageous and more realistic than social postwarnings because law officials may not always know who the witness has talked to and whether that source is reliable or not. Giving a sort of blanket statement about the potential harms of misremembering details
may prove to be more effective and pragmatic for police who are interpreting witness accounts or other forms of memory evaluations.

Echterhoff and Hirst (2005) examined the effects of different types of postwarnings on reports of misinformation. Specifically, they implemented the explicit postwarning condition and compared it against social postwarnings and no warnings. Using cued recall and recognition tests, they found that both social and explicit postwarnings diminished the misinformation effect significantly. While these results demonstrate the effectiveness of warnings, their methods for evaluating memory are not generalizable outside of the laboratory setting. Legal officials do not sanction simple cued recall and recognition tests, and they are almost never administered in police departments as methods for interviewing eyewitnesses. The present study explored a much more practical process to determine what participants remembered about a previously encountered event.

The cognitive interview (CI) was developed over 25 years ago as a tool for collecting and organizing eyewitness’ accounts (Fisher & Geiselman, 1992). Law enforcement officers and legal professionals alike frequently rely upon the cognitive interview as a method of memory retrieval (Memon, Meissner, & Fraser, 2010). The central idea behind the cognitive interview is to use research on what conditions are likely to produce accurate memories of past events as a basis for interviewing eyewitnesses. For example, one well-documented memory finding – referred to as the encoding-specificity principle – is that individuals are likely to optimally remember a past event when the conditions at retrieval match the conditions at encoding (Centofanti & Reece, 2006). Consequently, this principle is incorporated into the cognitive interview
REDUCING MISINFORMATION WITH WARNINGS AND COGNITIVE INTERVIEW

when an individual is asked to imagine and recreate the conditions in which the past event was experienced. This initial mental reinstatement phase, along with other phases in the interview, aim to alter the witness’s perspective and mode of thinking in terms of the crime (Memon et al., 2010). Another important psychological principle, known as the varied retrieval principle of memory, is integrated into the procedure of the cognitive interview (Centofani et al., 2006). Varied retrieval suggests that memory is a multifaceted combination of features instead of simply a holistic portrayal of the event, and, therefore, it is crucial to query memory in a variety of different ways so as to maximize the likelihood that the witness is recalling everything that she possibly can about the witnessed event (Centofanti et al., 2006). For instance, most cognitive interviews use a technique whereby the witness must recall the events of the crime in reverse chronological order, which is intended to increase overall rates of information recall by altering the perspective of the retrieval process.

We used an altered version of the modified cognitive interview (MCI) (Davis, McMahon, & Greenwood, 2005) that involved six distinct phases: build rapport, mental reinstatement, report everything, reverse order, tell me more, and confidence ratings. Newer versions of the cognitive interview recognized the importance of developing a comfortable, trusting relationship between the interviewer and the interviewee. This initial “Build rapport” phase is accomplished by engaging the interviewee in friendly conversation. Next, the interviewee is asked to mentally re-create the environment in which they witnessed the crime. “Report everything” is essentially a free recall memory test that asks the interviewee to divulge every possible detail no matter how seemingly insignificant about the event. The “Reverse order” phase requires the interviewee to
recount their memory for the event in reverse chronological order. During the “Tell me more” phase, the interviewer notes every reported detail and asks the interviewee whether or not they can say more about the event during which this detail occurred. Finally, the participant is asked to give confidence ratings for every detail that they mentioned. These ratings appraise the degree to which the interviewee is certain that they actually witnessed the reported item during the event.

As mentioned previously, misinformation can be introduced to a witness in a variety of different ways (Echterhoff et al., 2005). The present study used written narratives to suggest false information. Unlike previous studies that used questionnaires to implant these misleading details (Dodson & Krueger, 2006), we chose to use a narrative because it provided a more cohesive, relatable summary of the witnessed events. After watching a video depicting a robbery, participants read a narrative that mostly reiterated the actual chain of events but also casually mentioned certain misleading, incorrect information that seemingly fit into the schema of a typical crime, such as the robber wearing gloves. We wanted to see if postwarnings of misinformation could reduce the amount of misleading items recalled during the subsequent memory evaluation using the cognitive interview.

In addition to examining the amount of misleading items that participants recalled during the interview, we needed to see if the explicit warnings were also reducing the amount of correct information recalled. Therefore, a second group of items, known as tainted truth items, was planted into the narratives. Tainted truth items are accurate details that transpired in the video and were read about in the narrative. Nevertheless, they are “tainted” in the eyes of the participant because the warning serves to spoil the
participant’s perception of the narrative as a credible source (Echterhoff et al., 2007). Someone who remembers seeing and reading that the robber wore a denim jacket may be hesitant to report this item if they are warned about the possibility of false information. Echterhoff et al., (2007) found evidence for this tainted truth effect: warned participants recalled significantly fewer true items than unwarned participants, especially when the true items were integrated into the post-event source of misinformation (e.g. the narrative).

The cognitive interview has yet to be used in studies that employ explicit postwarnings to reduce the misinformation effect. We expect that if eyewitnesses are warned, thereby made to think suspiciously about their past memories, they will be less likely to report the misinformation on credible, applied memory tests like the cognitive interview. The cognitive interview should be particularly effective at determining whether the misinformation effect and the tainted truth effect are occurring because of its unique ability to draw out more detailed memories from witnesses during investigations (Memon et al., 2010). Moreover, we predict that the tainted truth effect will be diminished or eliminated in the warning condition because the cognitive interview focuses on reducing suggestibility and leading questions that might interfere with memory retrieval.

**Method**

**Participants**

Participants were 48 undergraduate students from the University of Virginia (10 men, 38 women, mean age= 18.60) who all gave consent prior to beginning the study. As
an incentive to participate, students received one and a half research-participation credit hours towards their required amount for the semester. Since the study involved some degree of deception, participants also signed a data release form after having completed the study in which they could choose to allow their data to be used or not.

**Design**

There were both between-participant and within-participant components in the design of this experiment. The between-participants component was whether participants were randomly assigned to either the warning condition or the no-warning condition. The within-participants component involved the type of items that we examined in order to assess the effect of the warning. In order to further analyze the potential benefits of the warning, we looked at the type of incorrect items reported (misleading vs. new) as a within-participants comparison. For analysis of harmful effects of the warning, we examined the “taintedness” of the items (tainted vs. untainted) as a second within-participants factor. The dependent measure, the cognitive interview, assessed what participants remembered (misleading items, true items, tainted true items, untainted true items, and new items) and their confidence for the accuracy of those memories.

**Materials and Procedure**

Participants were tested individually in private testing rooms and were randomly assigned to one of the two conditions (warning or no warning) and one of three versions of the study (A, B, or C) that served to counterbalance the items. After filling out a demographics form and signing the informed consent agreement, participants watched a
five-minute video of a crime on a 22-inch computer screen (in full screen mode). Following the video, there was a five-minute delay during which participants completed a filler task, such as a word search or crossword puzzle.

After the delay, participants read a post-event narrative (approximately 500 words) twice. To ensure that they were actually reading it completely through twice, participants read it once and then the experimenter asked them to read a second time only after they had indicated completing it the first time. The narrative served as a general summary of the events that transpired in the video, but it also included some false, or misleading, items. There were three versions of the narrative (A, B, or C) that counterbalanced the following critical items: 8 misleading items and 8 tainted truth items. Both sets of 8 were drawn from a larger pool of 12 whereby each version included a different group of 8 counterbalanced items. For example, the misleading item “the thief had a gun” is included in the A and C versions, but not the B version.

After reading the narrative, there was another ten-minute delay during which the participant is taken into a different testing room. This delay also marks the time when research assistants would administer two subtests – information and digit-symbol – from the Wechsler Adult Intelligence Scale (WAIS) and a visual acuity test.

The final portion of the pre-memory test segment of the study involved the receipt of a post-warning about the narrative. Depending on the condition, participants were given a piece of paper that said, “Please take a moment to think about the narrative that you read earlier” (no warning) or “Please be aware that the narrative you read earlier may have contained some misleading information. Please only report that which you remember happening in the video” (warning).
The next phase of the study focused on directly testing memory for the video using the cognitive interview. The cognitive interview itself has four distinct phases: mental reinstatement, reverse order, tell me more, and confidence ratings (Centofanti & Reece, 2006). An important aspect of the cognitive interview also involved building rapport with the participants, and this was achieved through casual conversation and questions prior to the cognitive interview. Experimenters would then conduct the interview by first asking participants to recall everything they remembered about the video (mental reinstatement). The experimenter used the cognitive interview scoring sheet as a guideline for recording the memories of the participants. Each memory item recalled was noted on the physical scoring sheet as well as the phase in which the item was recalled. Next, a reverse order phase was conducted whereby the participant was asked to recall everything starting at the end of the video and working backwards. A third phase known as “tell me more” involved the experimenter repeating every remembered detail back to the participant and asking them if there might be any more information they could remember about that particular detail. Finally, the researcher asked for confidence ratings for each and every item that was recalled. These ratings were scored on a 6-point scale from 50 to 100 in increments of 10.

Results

Every item recalled during the cognitive interview was categorized into one of six different source groups: video only, both (video and narrative), misleading, seen and read tainted truth, seen only tainted truth, and new. New items refer to anything mentioned that was completely false (i.e. not seen in the video) and not an experimentally implanted misleading item. Table 1 shows the frequency that these different items were reported by
individuals in the warning and no-warning groups. An independent samples t-test was conducted on each category separately, comparing the warning and no warning groups, but only the misleading category showed substantial results. There was a significant difference in the amount of misleading items recalled for the warning (M= .33, SD= .64) compared to the no warning (M=1.17, SD= 1.83) condition; t (46)= 2.10, p< .05. Participants were significantly less likely to report misleading details, suggested in the narrative, when they were warned about the possibility of encountering such misinformation (see Figure 1). There were no significant differences in the other five categories: all t (46)’s < 1.19.

In order to look more closely at the effects of the warning on the amount of tainted truth items recalled, we analyzed “tainted” and “untainted” items in a mixed ANOVA. Tainted items were items that had been both seen in the video and read about in the narrative, whereas untainted items were items that had only been seen in the video, not in the narrative, for a given participant. We computed a 2 (warning vs. no warning) x 2(item type: tainted vs. untainted truth) mixed ANOVA of the mean count of reported items between the two conditions. This particular test was run to determine if the warning produced any detrimental effects on participants’ memory, meaning anyone who reported fewer items as a result of the warning. There were no significant main effects or interactions, so we can conclude that warned participants did not filter out true items along with the false, or misleading, items (see Figure 2).

There were additional predictions about how the confidence ratings associated with each reported item might be affected by the warning. Would participants be less confident when warned about possible misinformation? To examine this question, we
conducted a 2 (condition: warning vs. no warning) x 3 (low, medium, or high confidence)
mixed ANOVA of the mean count of items recalled that were rated in varied confidence levels. The confidence ratings were grouped into three different ranges of high confidence (80-100), medium confidence (70-80), and low confidence (50-70). Although there were significantly more items rated with high confidence ($M = 13.35, SD = 13.14$) and medium confidence ($M = 13.58, SD = 13.42$) as compared to low confidence ($M = 0.83, SD = 1.41$), no significant differences were observed between the warning and no warning conditions.

**Discussion**

The most important message to take away from these results is that warnings can and do effectively reduce the misinformation effect amongst eyewitnesses. It is critical that persons involved in the judicial system understand psychological findings regarding eyewitness memory, and the effect of postwarnings on misinformation reduction is perhaps one of the most relevant findings to consider.

Echterhoff et al. (2005, 2007), Centofanti et al. (2006), and others have found similar results using warnings to reduce misinformation recall, but – until now – there has yet to be a study that measures the effectiveness of postwarnings using the more ecologically valid cognitive interview. Echterhoff et al., (2005, 2007) used basic laboratory memory tests, such as cued recall and recognition, to determine whether participants remembered false items, but these tests are not practically applicable for eyewitnesses and police investigators. The fact that the participants who were warned and then interviewed using the cognitive interview reported significantly fewer misleading details, as compared to the no-warning group, suggests that this combination of
procedures may be beneficial in the legal context. It might be possible to literally warn eyewitnesses about the influential nature of other sources that they may have encountered after the witnessed event. They can be made to understand misinformation and hopefully filter it from their memory prior to participation in the cognitive interview.

The tainted truth effect was not observed with the use of the cognitive interview, suggesting that witnesses were not restricting their responses or using a more stringent filter when reporting memories. They reliably included true items in their interviews even when the warning raised their suspicions about what is true and what is untrue. This finding is particularly relevant because it is contrary to what Echterhoff et al., (2007) observed in his experiments: that warned participants rejected accurate items from the post-event source to a greater extent than those who were not warned. It is possible that the varied phases of the cognitive interview allowed participants to think more cautiously about how and where they encoded their memories from the video and the narrative. This finding highlights another reason why the cognitive interview can better evaluate eyewitness memory.

This study was limited, in part, by the nature of the cognitive interview as the dependent measure. The cognitive interview is difficult to conduct consistently between participants. Despite having a team of research assistants who were well trained in the techniques of administering the interview (and having ten pilot studies prior to collecting data), interviewers inevitably interpreted recalled items in marginally different ways. This applied to both the way the interviews were conducted and the subsequent scoring of the data attained. The interview itself must be recorded by hand during the experiment, and the coding of this information can be fraught with complications and borderline
categorizations of reported items. Some reported memories simply did not differentiate clearly into one category. It would be best to have one person conduct and code all of the cognitive interviews, but this was simply not possible given the time constraints.

There were also some unexpected observations that came up during the study that may serve as opportunities to explore in future research. Most importantly, we found that many participants in the warning condition would recall misleading details in a semantically related manner. For example, one of the misleading items in the narrative was that the thief stole a gold ring from the house. Although this clearly did not occur in the video, many people would report seeing him grab a jewelry box or a necklace, which are also totally incorrect. We did not score these items as misleading unless they were specifically the same as what was written in the narrative (i.e. gold ring), but this effect is interesting nevertheless.

Future research should focus on alternative ways to reenact realistic eyewitnesses scenarios. Although the narrative proved to be an effective way to introduce the misinformation, it is not practical in the context of an authentic criminal event. Instead, future studies should include actors or additional experimenters to personally interact with participants while casually implying the false items in their dialogue. Similarly, warnings in these studies could be given as social warnings, which depict the post-event source as unreliable or untrustworthy.

Older adults (age 65 and above) are a unique and important group to study in terms of eyewitness memory. Based on past studies by Dodson & Krueger (2006), we know that older adults tend to make high confidence memory errors. Therefore, it is critical to determine how older adults might respond to the misinformation effect and the
cognitive interview. Given their history of high confidence mistakes, the warnings may not be successful at reducing recall rates of misleading information.
References


Table 1

*Frequency of Items Reported in Six Categories*

<table>
<thead>
<tr>
<th>Source</th>
<th>Warning (n=24)</th>
<th>No Warning (n=24)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Standard Error</td>
</tr>
<tr>
<td>Video</td>
<td>18.08</td>
<td>1.87</td>
</tr>
<tr>
<td>Both</td>
<td>19.17</td>
<td>1.29</td>
</tr>
<tr>
<td>Misleading</td>
<td>0.33</td>
<td>0.13</td>
</tr>
<tr>
<td>New</td>
<td>4.29</td>
<td>0.57</td>
</tr>
<tr>
<td>Total</td>
<td>41.88</td>
<td>1.93</td>
</tr>
<tr>
<td>Tainted Truth</td>
<td>3.08</td>
<td>0.31</td>
</tr>
<tr>
<td>Untainted Truth</td>
<td>1.00</td>
<td>0.21</td>
</tr>
</tbody>
</table>

*Note: Tainted and untainted items were scored separately from the video, both, misleading, and new categories. All tainted items were also counted in the both category because they are items that were both seen in the video and read in the narrative.*
Figure 1. Mean count data for amount of misleading items reported in the cognitive interview.
Figure 2. Mean count data for amount of tainted truth items reported in the cognitive interview.