Paper:
Horry, R., Colton, L. & Williamson, P. (2014). Confidence–accuracy resolution in the misinformation paradigm is influenced by the availability of source cues. *Acta Psychologica, 151*, 164-173.

http://dx.doi.org/10.1016/j.actpsy.2014.06.006

Confidence-Accuracy Resolution in the Misinformation Paradigm is Influenced by the Availability of Source Cues

Ruth Horry[a], Lisa-Marie Colton[a], and Paul Williamson[a1]

[a]Flinders University, Adelaide, Australia

Accepted for publication in *Acta Psychologica*. This article may not exactly replicate the final published version. It is not the copy of record.

Ruth Horry, School of Psychology, Flinders University, GPO Box 2100, Adelaide, South Australia, 5001, Australia. Email: Ruth.Horry@flinders.edu.au.

Lisa-Marie Colton, School of Psychology, Flinders University, GPO Box 2100, Adelaide, South Australia, 5001, Australia.

Paul Williamson, School of Psychology, Flinders University, GPO Box 2100, Adelaide, South Australia, 5001, Australia. Email: Paul.Williamson@flinders.edu.au. Phone: +61 8 8201 3644.

[1]Corresponding author

Abstract

After witnessing an event, people often report having seen details that were merely suggested to them. Evidence is mixed regarding how well participants can use confidence judgments to discriminate between their correct and misled memory reports. We tested the prediction that the confidence-accuracy relationship for misled details depends upon the availability of source cues at retrieval. In Experiment 1, participants (*N* = 77) viewed a videotaped staged crime before reading a misleading narrative. After seven minutes or one week, the participants completed a cued recall test for the details of the original event. Prior to completing the test, all participants were warned that the narrative contained misleading details to encourage source monitoring. The results showed that the strength of the confidence-accuracy relationship declined significantly over the delay. We interpret our results in the source monitoring framework. After an extended delay, fewer diagnostic source details were available to participants, increasing reliance on retrieval fluency as a basis for memory and metamemory decisions. We tested this interpretation in a second experiment, in which participants (*N* = 42) completed a source monitoring test instead of a cued recall test. We observed a large effect of retention interval on source monitoring, and no significant effect on item memory. This research emphasizes the importance of securing eyewitness statements as soon as possible after an event, when witnesses are most able to discriminate between information that was personally seen and information obtained from secondary sources.

Keywords: Misinformation effect, confidence, resolution, metacognition, source monitoring.

Confidence-Accuracy Resolution in the Misinformation Paradigm is Influenced by the Availability of Source Cues

1. Introduction

In a seminal study, Loftus et al. (1978) showed that people often incorporate misleading information encountered after a witnessed event into their memory reports of that event. Dozens of studies have replicated this *misinformation effect*, repeatedly showing that participants often report having seen details that were merely suggested (e.g., Belli et al., 1994; Chambers and Zaragoza, 2001; Lindsay, 1990). The metacognitive experiences associated with these errant memory reports have recently come under the empirical spotlight. A key question is whether participants are able to discriminate between their real and errant memory reports after exposure to misinformation. To date, evidence is mixed, with some researchers reporting very poor discrimination (Bonham and González-Vallejo, 2009; Cann and Katz, 2005; Tomes and Katz, 2000), and others reporting reasonably high discrimination (Higham et al., 2011). In this paper, we focus on the availability of source cues as a moderator of metacognitive discrimination in the misinformation paradigm. We argue that when source cues are relatively accessible, participants are better able to discriminate between their correct and suggested memories, but when source cues are relatively inaccessible, discrimination worsens.

1.1 Source monitoring and misinformation

The basic misinformation paradigm includes three stages. First, participants witness an event. The event is often depicted in slides (e.g., Frost et al., 2002; Higham et al., 2011; McCloskey and Zaragoza, 1989) or by video (e.g., Bonham and González-Vallejo, 2009; Cann and Katz, 2005), though live events have also been used, including those in which the participant was actively involved (e.g., Eisen et al., 2013; Holmes and Weaver, 2010; Sondhi and Gupta, 2007). Second, the participant is exposed to misinformation, which may be embedded in a narrative (e.g., Belli et al., 1992; Lindsay, 1990), incorporated into a series of post-event questions (e.g., Chambers and Zaragoza, 2001; Hekkanen and McEvoy, 2002), or presented by another witness (e.g., Meade and Roediger, 2002; Wright et al., 2000). Finally, the participant's memory is tested. Test formats have varied between studies, and have included recognition (e.g., Loftus et al., 1978; McCloskey and Zaragoza, 1985), cued recall (e.g., Belli et al., 1994; Thomas et al., 2010), free recall (e.g., Gabbert et al., 2006), and source memory tests (e.g., Lindsay and Johnson, 1989; Zaragoza and Koshmider, 1989). A consistent finding in these studies is that people often report having seen details that were merely suggested to them.

Although it was initially suggested that the original memory trace was irrevocably altered by the misinformation (Loftus et al., 1978; see also Greene et al., 1982; Loftus, 1979), there is now considerable evidence that the original memory trace can co-exist, unaltered, alongside the memory trace for the suggested detail (e.g., Christiaansen and Ochalek, 1983; Lindsay and Johnson, 1989). The upshot of this is that, under the right conditions, the original detail can be retrieved and the harmful influence of the misinformation can be undone (e.g., Gordon and Shapiro, 2012; Wright, 1993).

How might a participant resolve the discrepancy of having two conflicting memory traces available? According to the source monitoring framework (Johnson et al., 1993), when information is stored in memory, it is stored alongside various cues that can be used to infer the source of the information. These cues include perceptual details (e.g., visual and auditory details), spatial and temporal information, records of cognitive operations (e.g., elaboration, retrieval), and affective information. Consider a misinformation study in which a participant sees a hammer but later reads that the tool was a wrench. The memory for the hammer may be accompanied by perceptual details concerning its colour, shape, size, location within the scene, and so on. The memory for the wrench, however, may include perceptual details about the typeface in which the word was written, and the sound of the word as it was read. In a memory test, a participant could use these cues to discount the wrench and to correctly respond that the item was a hammer.

From the source monitoring perspective, performing accurately on a memory test following exposure to misinformation depends upon two factors. First, the participant must actively engage in source monitoring at retrieval. Second, diagnostic source cues (i.e., those that reliably differentiate between the event and post-event sources) must be available and accessible. Neither of these conditions, alone, will be sufficient for accurate performance. If a participant has ready access to diagnostic cues yet does not attempt to retrieve them, instead relying on retrieval fluency, then the participant will likely report some misinformation. Conversely, if a participant attempts to source monitor, but there are no (or very few) diagnostic cues available, source monitoring will be unsuccessful, potentially leading to reporting of misinformation. Below we present evidence that 1) participants do not automatically engage in source monitoring in misinformation tasks; and 2) even if participants are attempting to source monitor, the availability of source cues determines the likelihood of misinformation being reported.

First, there is considerable evidence that participants do not automatically engage in source monitoring. For example, test formats that encourage source monitoring typically produce smaller misinformation effects than testing conditions that promote responding based on retrieval fluency. Lindsay and Johnson (1989) showed participants a scene of a cluttered office before presenting them with a narrative containing several incorrect details. The participants were then presented with a list of items including event details and suggested details. Half of the participants made yes/no recognition decisions about whether each item had appeared in the picture, while the remaining participants made source judgments for each item. The proportion of items incorrectly attributed to the scene was .66 for the recognition test participants and only .32 for the source test participants (which was not significantly different from the control participants' error rate of .30). The authors argued that the recognition test participants had responded based on retrieval fluency, and had not engaged in source monitoring. The source test participants, however, could not rely on retrieval fluency, and so had to actively engage in source monitoring.

Further evidence that participants do not automatically engage in source monitoring comes from studies that have warned participants prior to the test that the post-event information contained incorrect information (these are often called *postwarnings*, as they are presented subsequent to the misinformation). Several studies have found that postwarnings reduce the size of the misinformation effect, at least under some conditions (e.g., Chambers and Zaragoza, 2001; Christiaansen and Ochalek, 1983; Echterhoff et al., 2005; Thomas et al., 2010). Because postwarnings are, by definition, presented after the misinformation has been encoded, their effectiveness cannot be due to differential encoding of the misleading details. Rather, the effects must be due to differences in retrieval processes. Specifically, it has been argued that warnings alert participants to the need to monitor the sources of their recollections. Given a postwarning, a participant who retrieves the suggested detail and who also recovers source cues that link the item to the post-event information may continue to search their memory for an alternative response. Without a postwarning, the participant may accept the suggested detail on the basis of its familiarity, thus terminating their memory search before the original detail is retrieved.

Even if a participant attempts to monitor the source of their memories, source misattribution errors will still occur. These errors should be relatively infrequent when diagnostic source cues are readily available, but common when source cues are unavailable. One factor that should have a large impact upon the availability of source cues is retention interval. Several studies have reported that the magnitude of the misinformation effect increases with longer retention intervals (e.g., Frost, 2000; Frost et al., 2002; Holmes and Weaver, 2010; Underwood and Pezdek, 1998). Frost et al. (2002), for example, showed participants a slide sequence in a first session, which was followed by a narrative containing some misleading details. Half of the participants completed a memory test 10 minutes later, while the remaining participants completed the memory test in a second session one week later. Participants in the one week delay condition were around 30-40% more likely to report misinformation than participants in the 10 minute delay condition. The authors concluded that there were fewer source cues available to participants after a longer delay, increasing the source similarity between the original and suggested details in memory.

In summary, source monitoring appears to play a central role in the production of misinformation errors. Participants will make fewer misinformation errors if they are encouraged to engage in source monitoring, but only if there are diagnostic source cues available at test.

1.2. Metacognitive monitoring and misinformation

Although the mechanisms that underlie the misinformation effect are still under debate, it is clear that participants often report misinformation. An important question, both theoretically and practically, is to what extent participants are able to discriminate between their correct and incorrect memories. A useful statistic for answering this question is *resolution*. Assessing resolution requires that participants respond to a reasonable number of items, assigning a confidence rating to each response. High resolution would be demonstrated if participants consistently assigned higher confidence ratings to their correct responses than to their incorrect responses; lower resolution would be demonstrated if there was considerable overlap in the confidence ratings for correct and incorrect responses.

A handful of studies have examined resolution after exposure to misinformation. Three studies have found reasonably strong resolution for control items but very poor resolution for misled items (Bonham and González-Vallejo, 2009; Cann and Katz, 2005; Tomes and Katz, 2000). In fact, in each of these studies, resolution for misled items was not significantly different from zero, indicating that the participants were unable to discriminate between their correct and misled responses. Tomes and Katz (2000) concluded that after the presentation of misinformation, “confidence becomes useless as an indicator of veracity” (p. 279). However, this conclusion may have been premature. In two experiments, Higham et al. (2011) reported similarly high resolution for misled items as for control items.

What could account for the discrepancy between the results of Higham et al. (2011) and those of prior studies? Higham et al. designed their procedure to encourage source monitoring by providing an explicit postwarning about the narrative. Participants were told that half of the questions on the test concerned details that had been incorrectly described in the post-event narrative, while the remaining questions concerned details that had not been described in the narrative. None of the other studies included a postwarning. In fact, in none were the participants even instructed to respond on the basis of what they remembered *seeing* rather than what they remembered *reading*. Higham et al. argued that participants in these studies would have had no reason to engage in source monitoring. They may well have assumed that the post-event information was a veridical source of information, and that any item that appeared familiar must therefore have been part of the original event (even if they could not explicitly recall seeing that item). Thus, engagement in source monitoring may be necessary for accurate discrimination between correct and misled responses.

Higham et al.’s (2011) results suggest that confidence judgments may be based on the same information as the memory decision itself. Without a postwarning (as in Bonham and González-Vallejo, 2009; Cann and Katz, 2005; Tomes and Katz, 2000), participants base their memory decisions and their confidence judgments on retrieval fluency, which can be misleading when misinformation has been presented (Lindsay and Johnson, 1989). This produces poor resolution for misled items. With a strong postwarning, participants base both their memory decisions and their confidence judgments on the available source cues (as in Higham et al.). Thus, when diagnostic source cues are readily available, participants can use confidence to discriminate between their correct and suggested memories.

Earlier we argued that the success of a source monitoring attempt will depend upon the availability of source cues at retrieval and that source cues degrade over an extended retention interval. If participants base their confidence judgments on the same source cues that are used to produce the memory decision, then we might predict that longer retention intervals should reduce resolution for misled items. The aim of this study is to provide further evidence for the source monitoring hypothesis of confidence judgments, by examining the impact of retention interval on resolution.

1.3. The present study

Four studies have examined resolution within the misinformation paradigm (Bonham and González-Vallejo, 2009; Cann and Katz, 2005; Higham et al., 2011; Tomes and Katz, 2000). In each of these experiments, the event, the misinformation, and the memory test took place within a single session. Thus, the retention intervals were mere minutes. In the present study, we showed our participants an event and presented them with a misleading narrative in a first session. Half of the participants also completed a memory test in the first session; the remaining participants returned one week later to complete the memory test. Before the memory test, all participants were warned that the narrative contained inaccurate details. In the interests of generalizability, we chose to use a cued recall test in the present study. This method contrasts with the yes/no recognition tests (Bonham and González-Vallejo, 2009; Cann and Katz, 2000; Tomes and Katz, 2000) and *n*-alternative forced choice tests (Higham et al., 2011) used previously. Thus, a conceptual replication of earlier findings with an alternative testing method would increase our confidence in the reliability of our findings.

Our primary hypothesis was that resolution for misled items would decline substantially over the retention interval, and at a higher rate than resolution for control items. Consequently, we expected that we would replicate the results of Higham et al. (2011) in the short delay condition, with good resolution for control and misled items. In the long delay condition, however, we predicted that we would find good resolution for control items but poor resolution for misled items, in line with previous studies (Bonham and González-Vallejo, 2009; Cann and Katz, 2005; Tomes and Katz, 2000).

In addition to confidence judgments, we also asked participants whether they would like to testify or withhold each response (see Higham et al., 2011). Type-2 signal detection measures were then calculated to provide a behavioural index of participants' abilities to discriminate between their own correct and incorrect answers (Higham et al., 2009). We expected these data to complement the confidence judgments. Specifically, we expected good discrimination for misinformed items in the short delay condition but poor discrimination in the long delay condition.

2. Experiment 1

2.1. Method

2.1.1. Participants and Design

Eighty participants took part for credit on an introductory Psychology course or for payment. Two participants were excluded as they were not fluent in English; all remaining participants were native English speakers. One participant was excluded for failing to follow instructions, leaving 77 participants in the final analyses. Fifty four participants were female, 22 were male, and one did not provide his or her gender. The mean age was 22.45 years (*SD* = 9.12 years).

The experiment followed a 2 (Item Type: control, misled) × 2 (Retention Interval: immediate, delayed) mixed design, with repeated measures on the first factor. Thirty seven participants were randomly allocated to the immediate testing condition and 40 were randomly allocated to the delayed testing condition.

2.1.2. Materials

The event was a videotaped staged crime, in which a thief, posing as a gas installation worker, robs several items from an elderly man's home. The event was approximately four minutes long.

Pilot testing was undertaken to ensure that 1) memory for each of the critical items was reasonable, even after a one week delay; and 2) the misleading suggestions were all seen as plausible. In the first pilot test, participants (*N* = 15) watched the video and then completed a cued recall questionnaire immediately and again after one week. Only items that were remembered by more than 50% of participants were selected as potential critical items. In the second pilot test, a new group of participants (*N* = 15) watched the video and then completed a 42-item questionnaire. Each question referred to a detail from the video, and had four possible responses. One response was the correct detail, while the remaining responses were incorrect. Participants were asked to choose which alternative had actually been presented in the video and were asked to rate the plausibility of each other alternative on a 1 (high implausible) to 7 (highly plausible) scale. Items with mean plausibility ratings of 4-5 were selected as the misleading details for the study. After pilot testing, 32 items were selected as critical details.

The post-event information was presented in a 480 word narrative (see Appendix). Two versions of the narrative were created, each including 16 incorrect details. All of these details contradicted a detail that was seen in the video. For example, the video showed the old man sitting in a wooden rocking chair, but the relevant misleading detail was that he was sitting in a leather chair. The 16 critical details that served as misleading items for half of the participants served as control items for the remaining participants.

Most of the control items were not mentioned in the narrative, or were mentioned only in a neutral form. For example, one critical item concerned the colour of the thief's jumper (navy blue). In one version of the narrative, participants read that the thief was "dressed in a *green* jumper" (misleading detail); in the other version, the thief is simply described as "dressed in a jumper" (control detail). However, to ensure that the narrative flowed coherently, four control details were mentioned in their correct, specific form in each version of the narrative. For example, one narrative mentioned that the thief offered to provide a free quotation on gas installation to gain entry to the victim's house.

The memory test consisted of 32 cued recall questions (16 relating to the misinformation items and 16 to the control items), each corresponding to one of the critical details. The questions were worded to elicit a fine-grained response: for example, "What kind of chair was the old man sitting in?" Next to the space for each answer was a confidence scale, showing numbers from 0% to 100% in 10% increments.

2.1.3. Procedure

Participants signed up for a study on "the effect of delayed recall on memory for event details". Participants were tested individually or in groups of two, though they worked independently at their own computer terminal throughout the whole experiment. Upon arrival at the laboratory, participants watched the staged crime video. They were told to pay close attention to the video but they were not explicitly informed that they would have their memories tested on any specific aspects of the event.

Following the video, participants completed a filler task (a series of mazes) for seven minutes. All participants were then presented with one of the two versions of the post-event narrative. To ensure that the information was being processed, the narrative was presented on six pieces of card, in a scrambled order. Participants were asked to read each card carefully and to sort them into chronological order, and the task was presented under the guise of a "comprehension task" (see Higham et al., 2011; Zaragoza and Lane, 1994).

Following the narrative task, participants completed a second filler task (additional mazes) for seven minutes. Participants who had been assigned to the immediate condition proceeded to the memory task, whereas participants in the delayed condition were dismissed. These participants returned one week later to complete the memory test. Crucially, participants did not know which delay condition they had been assigned to until this point in the experiment. Thus, there was no possibility that participants differentially encoded the information in anticipation of a short or long delay.

Participants received the following instructions prior to the memory test: "Please think back to the film you watched, and for every question write down what you remember seeing in the film. Please also rate how confident you are that each response you give is correct. It is important that you report only what you remember seeing, because the narrative you read contained some incorrect details. If you really cannot remember, please make your best guess (i.e., write down the first thing that comes to mind)".[1]

After completing the memory test, participants were asked to look back through their responses and to consider whether they would testify to each response. Specifically, participants were given the following instructions: "We now want you to imagine that the questions you have just been asked in the questionnaire have been asked in a court of law, and the responses you give will be considered as your eyewitness testimony. You are the only eyewitness of the robbery, therefore your testimony needs to be as accurate as possible, as any incorrect responses could have serious consequences for this investigation. With this in mind, please go through and write a "T" for testify in the left-hand box for the answers that you would be willing to testify in a court of law. For those responses you would not offer as evidence in a court of law, write "W" for withhold." Thus, we used a "two-pass" procedure, in which participants made their testify/withhold decisions on a second lap of the questionnaire. The two-pass procedure has been criticized (e.g., Higham and Arnold, 2007) as the extra processing time afforded by the second lap may cause participants to change their minds (for example, if they recalled additional information between making the response and making the testify/withhold judgment). However, as our main concern was with resolution, we chose to use the two-pass procedure to avoid any possibility that the testify/withhold judgments were contaminating the confidence judgments.

After all of the tasks were completed, participants were thanked, debriefed, and asked not to discuss the experiment with anyone else.

2.2. Results

2.2.1. Coding

Each response was coded as correct, incorrect-guess, or incorrect-misinformation. Where responses incorporated some aspect of both the original detail and the misleading detail (i.e., a ‘memory blend’; Belli, 1988; Skagerberg and Wright, 2008a), the response was coded as incorrect-misinformation. Two blinded independent raters coded the responses. Agreement was 97.5%, and disagreements were resolved by discussion.

2.2.2. Response accuracy

Table 1 shows the proportion of correct, incorrect-guess, and incorrect-misinformation responses to control and misleading details. Note that no misleading details were ever reported for control items (i.e., no participant confabulated a detail that was presented as a misleading suggestion in the alternative narrative). Separate 2 (Item Type) x 2 (Retention Interval) mixed ANOVAs were conducted for the correct responses and incorrect-guesses. For the incorrect-misinformed responses, an independent samples *t* test compared response rates across the immediate and delayed conditions. For all pairwise comparisons, Cohen’s *d* is reported as a measure of effect size, using formulae appropriate for repeated measures contrasts or between groups contrasts as appropriate (see Dunlap et al., 1996). For interaction terms, partial eta square (η_p^2) is reported as a measure of effect size.

For correct responses, both main effects were significant. Accuracy was higher for control items than for misleading items, $F(1, 75) = 28.77$, $p < .001$, $d = 0.73$, and after a short delay than after a long delay, $F(1, 75) = 31.37$, $p < .001$, $d = 1.29$. The interaction term was not significant, $F(1, 75) = 0.01$, $p = .99$, $\eta_p^2 < .01$.

For incorrect-guesses, both main effects were significant. Guessing rates were higher for control items than for misinformation items, $F(1, 75) = 43.46$, $p < .001$, $d = 0.93$. Guessing rates also increased over the delay period, $F(1, 75) = 21.28$, $p < .001$, $d = 1.07$. The interaction term was not significant, $F(1, 75) = 1.19$, $p = .28$, $\eta_p^2 = .02$.

Finally, misinformation was no more likely to be reported after a long delay than after a short delay, $t(75) = 1.26$, $p = .21$, $d = 0.29$.

2.2.3. Resolution

To allow for direct comparisons with prior research (e.g., Bonham and González-Vallejo, 2009; Cann and Katz, 2005; Higham et al., 2011; Tomes and Katz, 2000), we computed Goodman-Kruskal gamma correlations (*G*) for control items and misled items separately, which is a measure of association appropriate for variables on ordinal scales (Nelson, 1984). Scores can range from -1 (perfect *negative* discrimination: all inaccurate responses assigned higher confidence ratings than all accurate responses) to +1 (perfect *positive* discrimination: all accurate responses assigned higher confidence ratings than all inaccurate responses). For control items, *G* was computed using correct and incorrect-guess responses; for misleading items, *G* was calculated using correct and incorrect-misinformation responses (i.e., incorrect-guess responses were not included in the *G* coefficient for misled items). As in Higham et al. (2011), we included both testified and withheld responses in the resolution scores. Two participants from the immediate condition and two from the delayed condition were excluded from the analyses as they did not report any misinformation.

The mean *G* coefficients are show in Figure 1. The coefficients were compared in a 2 (Item Type) x 2 (Retention Interval) mixed ANOVA. Both main effects were significant. Resolution was higher for control items (M = .74, SD = .19) than for misinformation items (M = .46, SD = .40), $F(1, 71) = 37.29$, $p < .001$, $d = 0.93$. Resolution was also higher in the immediate condition (M = .72, SD = .17) than in the delayed condition (M = .50, SD = .23), $F(1, 71) = 24.64$, $p < .001$, $d = 1.16$. However, these main effects should be interpreted in the light of a significant interaction term, $F(1, 71) = 10.05$, $p = .002$, $\eta_p^2 = .12$.

Our primary hypothesis was that retention interval would moderate resolution for misled items. An independent samples *t* test confirmed that retention interval had a large effect on resolution for misled items, $t(71) = 4.61$, $p < .001$, $d = 1.08$. Resolution was substantially higher in the immediate condition (M = .65, SD = .24) than in the delayed condition (M = .27, SD = .43). Retention interval had a smaller effect on resolution for control items, $t(71) = 1.84$, $p = .07$, $d = 0.43$, with marginally higher resolution in the immediate condition (M = .79, SD = .22) than in the delayed condition (M = .71, SD = .15).

For comparability with prior studies, we also compared the mean resolution scores for control items and misled items in each condition. In the delayed condition, the control-misinformation difference was large, $t(37) = 5.78$, $p < .001$, $d = 1.34$. The control-misinformation difference was also significant in the immediate condition, though the effect was smaller, $t(34) = 2.58$, $p = .01$, $d = 0.59$.

One potential criticism of the above analyses is that the resolution coefficients for control and misled items are based on different information (correct *vs.* incorrect-guesses, and correct *vs.* incorrect-misinformation, respectively).Arguably, the control coefficient measures reality monitoring (the extent to which one can distinguish between real and confabulated memories; Johnson and Raye, 1981) while the misled coefficient measures source monitoring (the extent to which one can distinguish between two external sources of information). To allay these concerns, we recalculated the gamma correlations for the misled items to include all erroneous responses. The results were similar to those reported above: resolution was higher in the immediate condition (.67, *SD* = .23) than in the delayed condition (.39, *SD* = .32), $t(72) = 4.20$, $p < .001$, $d = 0.98$. Thus, our conclusions hold even if guesses are included in the resolution index for misled items, though the effect size is somewhat reduced.

Another potential criticism is that some of the control details were included in the narrative, while the majority were not. We recalculated the gamma coefficients for the control items, including only the items that were not in the narrative. One additional participant was excluded from this analysis as she made no errors in the subset of control items included. The means changed very little, leading to a very similar pattern of results. Specifically, the main effects of item type, $F(1, 70) = 27.37$, $p < .001$, $d = 0.83$, and retention interval, $F(1, 70) = 20.51$, $p < .001$, $d = 1.08$, were significant, as was the crucial interaction term, $F(1, 70) = 7.58$, $p = .008$, $\eta_p^2 = .10$.

To ensure that these results were not simply a product of reduced recall accuracy in the delayed condition (Perfect and Stollery, 1993), we ran a (multi-level) linear model analysis, in which recall accuracy was entered as a covariate. The interaction between item type and recall condition remained significant, $F(1, 114.55) = 9.77$, $p = .002$, indicating that the results cannot be explained by the reduction in recall accuracy in the delayed condition.

2.2.4. Type-2 Signal Detection

Type-2 d' (see Higham, 2002) was calculated for control items and misinformed items. Unlike the correlation coefficient, d' is unbounded. Negative values indicate negative discrimination (more incorrect answers testified than correct answers), and positive values indicate positive discrimination (more correct answers testified than incorrect answers). For both control and misled items, the hit rate was defined as the proportion of correct responses that were testified. For control items, the false alarm rate was defined as the proportion of incorrect-guesses that were testified; for misled items, the false alarm rate was defined as the proportion of incorrect-misinformation responses that were testified. For example, if a participant reported six misled details and chose to testify three of them, their false alarm rate would be .50. Again, two participants from the immediate condition and one from the delayed condition were excluded from the analyses as they did not report any misinformation.

Mean Type-2 *d′* values are shown in Figure 2. It is immediately apparent that the pattern of means closely resembles the pattern for the resolution coefficients in Figure 1. A 2 (Item Type) x 2 (Retention Interval) mixed ANOVA revealed significant main effects of Item Type, $F(1, 71) = 29.46$, $p < .001$, $d = 0.85$, and Retention Interval, $F(1, 71) = 7.55$, $p = .008$, $d = 0.65$. Similar to the earlier gamma correlation findings, these main effects need to be interpreted in the light of a significant interaction term, $F(1, 71) = 9.39$, $p = .003$, $\eta_p^2 = .12$.

In support of our primary hypothesis, retention interval had a large impact on discriminability of misled items, $t(71) = 3.52$, $p = .001$, $d = 0.84$, with significantly better discrimination in the immediate condition ($M = 1.00$, $SD = 0.66$) than in the delayed condition ($M = 0.40$, $SD = 0.79$). In contrast, retention interval did not significantly affect discrimination of control items, $t(72) = 0.10$, $p = .92$, $d = 0.02$.

We also compared discriminability for control and misled details. In line with Higham et al.’s (2011) findings, we found no significant difference in the immediate condition, $t(71) = 1.76$, $p = .09$, $d = 0.40$. However, in the delayed condition, the control-misled difference was large, $t(71) = 5.79$, $p < .001$, $d = 1.31$.

Once again, when we controlled for recall accuracy in a multi-level linear model analysis, the interaction term remained significant, $F(1, 128.30) = 9.30$, $p = .003$.

2.2.5. Accuracy and Resolution for Testified Responses

Given the option, participants tend to withhold low confidence answers. The result is almost always an increase in accuracy, as low confidence answers are often erroneous (Koriat and Goldsmith, 1996). Our results showed that participants were able to increase their accuracy by choosing which answers to withhold. For control items, proportion correct increased from .71 (*SD* = .12) to .95 (*SD* = .37), $t(76) = 5.66$, $p < .001$, $d = 0.88$. For misled items, proportion correct increased from .61 (*SD* = .16) to .81 (*SD*= .20), $t(76) = 11.88$, $p < .001$, $d = 1.11$. The proportion of misinformation that was testified (.18, *SD* = .19) was significantly lower than the proportion that was reported in the forced choice task (.24, *SD* = .15), $t(76) = 3.27$, $p = .002$, $d = 0.34$, indicating that participants were successful in withholding misinformation at least some of the time.

An interesting question is whether providing participants with the option to withhold responses affects resolution. We recalculated the gamma coefficients to include only testified responses. As these coefficients require at least one incorrect testified response, we could only calculate them for a subset of participants. The *G* for testified misled items included 49 participants and the *G* for testified control items included 32 participants. We compared resolution for the total set of items with resolution for testified items in two paired-samples *t* tests. Resolution was not affected by withholding responses for either the control items (all responses: $M = .70$, $SD = .19$; testified responses: $M = .65$, $SD = .48$), $t(31) = 0.71$, $p = .48$, $d = 0.15$, or the misled items (all responses: $M = .38$, $SD = .36$; testified responses: $M = .23$, $SD = .63$), $t(48) = 1.70$, $p = .10$, $d = 0.30$. We note, however, that the gamma coefficients for the testified responses showed much greater variability than the corresponding coefficients for the full response set.

2.3. Discussion

The results of Experiment 1 supported our main hypothesis that resolution for misleading details would significantly decline over a delay. In the immediate condition, the confidence-accuracy relationship for misleading details was reasonably strong (as in Higham et al., 2011), though less strong than for control details. After one week, however, the confidence-accuracy relationship was much weaker – though significantly greater than the zero relationship reported in some prior studies (Bonham and González-Vallejo, 2009; Cann and Katz, 2005; Tomes and Katz, 2000). We believe that the most parsimonious explanation for these results is that source cues degrade rapidly over time, reducing the accessibility of diagnostic source cues in the delayed condition. In the absence of available source cues, participants would have been forced to rely on familiarity, which is an invalid accuracy cue after exposure to misinformation.

However, it is possible that other processes, such as item memory, were affected by the delay. If item memory were to decay at different rates for control and misleading details, the relative accessibility of the details at test might have downstream effects on confidence and, therefore, resolution. In Experiment 2 we directly tested our assumptions that retention interval affected source memory to a greater extent than item memory, and that any change in item memory over time was similar for control and misleading details. The method was identical to Experiment 1 except that the cued recall test was replaced by a source monitoring test (Lindsay & Johnson, 1989).

3. Experiment 2

3.1. Method

3.1.1. Participants and Design

A new sample of 42 undergraduate participants took part for credit on an introductory psychology course. Participants were randomly allocated to the immediate (n = 20) or delayed (n = 22) condition. The mean age was 23.11 years (SD = 7.60 years); 26 (62%) of the participants were female.

The experiment followed a 2 (Item Type: control, misled) × 2 (Retention Interval: immediate, delayed) mixed design, with repeated measures on the first factor.

3.1.2. Materials

The video, narratives, and filler tasks were identical to those used in Experiment 1. The source memory test included 32 items, each corresponding to one of the critical details from Experiment 1. For each item, the participant was presented with a statement (e.g., “The old man was sitting in a wooden rocking chair”), and four response options: “FILM only”, “TEXT only”, “BOTH”, or “NEITHER” (Lindsay & Johnson, 1989). A space was also provided for participants to circle their confidence from 0% to 100% in 10% increments. Two versions of the test were created, one for each narrative version. Participants were questioned about misleading details and the control details that corresponded to the narrative that they had read; they were not questioned about misleading details from the other narrative. So, for example, Test A items included the 16 critical misleading details from narrative A, the 12 critical control details mentioned only in neutral or superordinate form in narrative A, and the 4 consistent details mentioned in narrative A. The consistent details were excluded from the analyses, as the correct source attribution for these items would be “both”.

3.1.3. Procedure

The procedure was identical to Experiment 2, with the exception of the change in test format. Before completing the source memory test, participants were warned “It is important that you say that an item appeared in the film only if you remember seeing it, because the narrative contained some incorrect details”. They were told to indicate, for each statement, whether the detail was a) present only in the film; b) present only in the text; c) present in both the film and the text; or d) present in neither the film nor the text. They were also told to indicate their confidence in each decision by circling one of the 11 available confidence ratings[2].

3.2. Results

Table 2 shows the proportion of video, text, both, and neither responses by item type and delay. For each of these four response types, independent samples *t* tests were run to compare the mean proportions of responses in the immediate and delayed conditions, separately for control and misleading items. To control family-wise error, we used a Bonferroni-corrected α of .006 (.05/8). Only one comparison was significant: The proportion of misleading items correctly attributed to the text was significantly higher in the immediate condition than in the delayed condition, $t(38) = 4.10$, $p < .001$, $d = 1.33$.

To estimate item memory, we summed video, text, and both responses. The item memory parameters are shown in Table 2. We analyzed item memory in a 2 (Item Type) × 2 (Delay) mixed ANOVA. The main effect of item type was significant, $F(1, 39) = 13.52$, $p = .001$, $d = 0.73$, indicating that item memory was higher for control details ($M = .77$, $SD = .15$) than for misleading details ($M = .66$, $SD = .15$). However, neither the main effect of delay, $F(1, 39) = 0.09$, $p = .76$, $d = 0.10$, nor the Item type × Delay interaction, $F(1, 39) = 1.57$, $p = .22$, $\eta_p^2 = .04$, were significant. Thus, we observed no significant decline in item memory over the one week retention interval.

To estimate source memory, we calculated the proportion of recognized items that were attributed to the correct source (e.g., for control items: Film responses/(Film responses + Text responses + Both responses)). The source memory parameters are shown in Table 2. We analyzed source memory in a 2 (Item Type) × 2 (Delay) mixed ANOVA. The main effect of item type was significant, $F(1, 39) = 21.34$, $p < .001$, $d = 0.90$, indicating that source memory was higher for control details ($M = .59$, $SD = .25$) than for misleading details ($M = .37$, $SD = .24$). Importantly, the main effect of delay was also significant, $F(1, 39) = 15.43$, $p < .001$, $d = 1.26$, as source memory was higher in the immediate condition ($M = .58$, $SD = .18$) than in the delayed condition ($M = .38$, $SD = .14$). The Item type × Delay interaction was not significant, $F(1, 39) = 1.35$, $p = .25$, $\eta_p^2 = .03$. Thus, in contrast to item memory, source memory significantly declined over the retention interval, and similarly for control and misleading items.

3.3. Discussion

The results of Experiment 2 suggest that the assumptions underlying our interpretation of the results from Experiment 1 were sound. Specifically, source memory declined significantly over the retention interval, and the effect was large. Given that an item was remembered, there was a 58% chance that it would be attributed to the correct source in the immediate condition, but that probability dropped to 38% after a one-week delay. In contrast, we observed no significant decline in item memory over the delay. In conclusion, and in line with prior research (e.g., Frost, 2000; Frost et al., 2002; Holmes and Weaver, 2010; Underwood and Pezdek, 1998), source memory appears to degrade more rapidly than item memory.

4. General Discussion

We tested the prediction that, when participants are encouraged to engage in source monitoring, the confidence-accuracy relationship for misled details would be moderated by the availability of source cues at retrieval. As a manipulation of source cue availability, we varied retention interval, testing participants' memories either seven minutes or one week after exposure to the misinformation. Experiment 2 confirmed the validity of this manipulation, as it showed that retention interval had a large effect on source monitoring, but no significant effect on item memory. In Experiment 1, our primary hypothesis was supported; retention interval had a large impact on two measures of metacognitive discrimination for misled details: resolution and Type-2 d' (Higham, 2002). In addition, we hypothesized that, when source cues were readily available, resolution would be high for control items *and* misled items (as in Higham et al., 2011). In contrast, when source cues were relatively inaccessible, we expected to find a large difference in resolution for the control and misled items (as in Bonham and González-Vallejo, 2009; Cann and Katz, 2005; Tomes and Katz, 2000). These predictions were also supported.

We argue that the source monitoring framework (Johnson et al., 1993) provides the simplest explanation of our findings. Figure 3 illustrates the decision pathways that might be involved in a cued recall task in which the participant is actively engaged in source monitoring. First, the participant generates candidate answer(s) in response to a specific question. If multiple candidates are generated, the participant selects the best candidate for closer scrutiny. The participant then attempts to retrieve cues that will help to diagnose the source of the remembered detail. If event-related cues are retrieved (e.g., visuo-spatial details), then the participant reports the answer with high confidence. If post-event cues are retrieved (e.g., a memory of the typeface in which the word was written), the participant faces a decision about whether to report the item or return to their candidate answers to evaluate the next-best candidate. Should the participant decide to report the answer, he should assign low confidence to the reported detail (as he is aware that it came from the post-event source). Finally, if no source cues can be retrieved, the participant may return to his candidate answers or he may proceed to an assessment of retrieval fluency. If the retrieval fluency is below some pre-determined decision threshold, the participant may return to his candidate answers rather than report that answer. However, if the retrieval fluency exceeds the threshold, the participant will likely report the item, despite a lack of source cues. The confidence rating assigned to the memory report will be scaled to the retrieval fluency (Norman and Wickelgren, 1969), which may be misleading in the context of a misinformation experiment (Lindsay and Johnson, 1989).

According to this model, there are two reasons why a participant might report a suggested detail with high confidence: if event-related cues are incorrectly retrieved in association with the suggested detail (a *source misattribution error*); or if the detail was lacking in source cues but was associated with high retrieval fluency (a *fluency error*). Both error types may increase over time as source cues degrade (e.g., Frost, 2000; Frost et al., 2002; Holmes and Weaver, 2010; Underwood and Pezdek, 1998). If few source cues are accessible, participants may rely more heavily upon fluency, increasing fluency errors. Participants may also adjust their source monitoring criteria over time, requiring less evidence to attribute a recalled detail to the event, thus increasing source misattribution errors. Similar criterion shifts have been reported in old/new recognition tasks. Strong memories lead participants to set strict decision criteria, while weaker memories encourage more lenient criteria (Stretch and Wixted, 1998). Disentangling source misattribution errors from fluency errors may well be difficult, as they may often co-occur. To isolate their relative contributions, future research could use procedures that strongly discourage reliance upon retrieval fluency at test (e.g., Lindsay, 1990; Wright, 1993).

One potential avenue of research is to investigate resolution under conditions that should increase source misattribution errors. To our knowledge, the only study to touch on this issue is Higham et al. (2011). In Experiment 1, the authors omitted the control details from the narrative. In Experiment 2, however, they included the control details in the narrative, reasoning that source monitoring would be more difficult as post-event source cues (e.g., memory for the typeface of the word) would not reliably discriminate between control and misled items. However, resolution was similar across the two experiments, indicating that the change in method had little impact. We also included some control items in the narrative, yet our results changed little whether or not these items were included in the resolution scores. Perhaps other manipulations will prove more powerful, such as increasing the similarity of the event and post-event sources (e.g., Lindsay, 1990), or encouraging the use of mental imagery during the encoding of the post-event information (see related research on imagination inflation: Garry et al., 1996; Heaps and Nash, 1999; Thomas et al., 2003).

In Experiment 1, we allowed our participants to withhold any of their forced-choice answers in a subsequent free report phase. In line with previous research, participants increased their accuracy by screening out low-confidence errors (Koriat and Goldsmith, 1996). Resolution was not significantly affected by the change from forced to free report. However, this null effect should be interpreted cautiously. The resolution estimates for the testified responses were unstable due to the low number of errors reported; consequently, the error variance was large. Furthermore, Perfect and Weber (2012) showed that the order of free- and forced- responding influences patterns of final responses in memory tasks. Thus, a between-subjects manipulation seems necessary to resolve the question of whether resolution is influenced by free versus forced report conditions.

A limitation of this research is that retention interval is likely to influence factors other than source memory. Experiment 2 allowed us to rule out one obvious candidate, item memory, yet there are other potential candidates, including guessing biases, retrieval fluency, and meta-cognitive beliefs. It is possible, therefore, that some combination of these factors may have contributed to our results. A potential avenue for future research is to explore the influence of more “process-pure” manipulations of source memory, such as a manipulation of instructions regarding the presence of misinformation, on resolution. Nonetheless, the most parsimonious explanation of the present findings is that the accessibility of source cues degraded over the retention interval, which in turn disrupted participants’ abilities to monitor the accuracy of their reports. However, we must acknowledge that there may have been additional factors contributing to the effect.

From a practical perspective, our results underscore the importance of securing eyewitness statements as quickly as possible after an event has taken place. Not only will the opportunities for exposure to misinformation be reduced, but participants are able to discriminate between seen and suggested details, if encouraged to do so. Furthermore, there is some evidence that eliciting detailed free recall statements soon after the event reduces forgetting rates and protects witnesses from the harmful effects of misinformation that is later encountered (Gabbert et al., 2012). Even in cases with limited police resources and multiple witnesses (the sorts of situations in which co-witness transmission of misinformation is likely to occur; Skagerberg and Wright, 2008b), tools now exist to allow witnesses to record their own detailed statements (the Self-Administered Interview; Gabbert et al., 2009). Taken together with the results of Higham et al. (2011), our results suggest that the most effective way to minimise the harm of misinformation is to secure detailed witness statements at the earliest possible point and to encourage source monitoring at retrieval.

5. Conclusions

The strength of the confidence-accuracy relationship following exposure to misinformation depends crucially upon: 1) active engagement in source monitoring; and 2) the availability of diagnostic source monitoring cues at retrieval. If source cues are easily acessible, and if participants are encouraged to retrieve them, participants are able to use confidence judgments and testify/withhold judgments to discriminate between real and suggested memories. However, if diagnostic source cues are relatively inaccessible (for example, due to an extended delay between encoding and retrieval), participants find it much more difficult to discriminate between their real and suggested memories.

References

Belli, R.F., 1988. Colour blend retrievals: Compromise memories or deliberate compromise responses? Mem Cognit, 16, 314-326.

Belli, R.F., Windschitl, P.D., McCarthy, T.T., Winfrey, S.E., 1992. Detecting memory impairment with a modified test procedure: Manipulating retention interval with centrally presented event items. J Exp Psychol Learn Mem Cogn, 18, 356-367.

Belli, R.F., Lindsay, D.S., Gales, M. ., McCarthy, T.T., 1994. Memory impairment and source misattribution in postevent misinformation experiments with short retention intervals. Mem Cognit, 22, 40-54.

Bonham, A,J., González-Vallejo, C., 2009. Assessment of calibration for reconstructed eye-witness memories. Acta Psychol, 131, 34-52.

Cann, D.R., Katz, A.N., 2005. Habitual acceptance of misinformation: Examination of individual differences and source attributions. Mem Cognit, 33, 405-417.

Chambers, K.L., Zaragoza, M.S., 2001. Intended and unintended effects of explicit warnings on eyewitness suggestibility: Evidence from source identification tests. Mem Cognit, 29, 1120-1129.

Christiaansen, R.E., Ochalek, K., 1983. Editing misleading information from memory: Evidence for the coexistence of original and postevent misinformation. Mem Cognit, 11, 467-475.

Dunlap, W.P., Cortina, J.M., Vaslow, J.B., Burke, M.J., 1996. Meta-analysis of experiments with matched groups or repeated measures designs. Psychol Methods, 1, 170-177.

Echterhoff, G., Hirst, W., Hussy, W., 2005. How eyewitnesses resist misinformation: Social postwarnings and the monitoring of memory characteristics. Mem Cognit, 33, 770-782.

Eisen, M.L., Gomes, D.M., Lorber, W.G., Perez, C.I., Uchishiba, H., 2013. Using an individual differences approach to examine two distinct types of suggestibility effects. Appl Cogn Psychol, 27, 2-11.

Frost, P., 2000. The quality of false memory over time: Is memory for misinformation "remembered" or "known"? Psychon Bull Rev, 7, 531-536.

Frost, P., Ingraham, M., Wilson, B., 2002. Why misinformation is more likely to be recognised over time: A source monitoring account. Memory, 10, 179-185.

Gabbert, F., Hope, L., Fisher, R.P., 2009. Protecting eyewitness evidence: Examining the efficacy of a self-administered interview tool. Law Hum Behav, 33, 298-307.

Gabbert, F., Hope, L., Fisher, R.P., Jamieson, K., 2012. Protecting against misleading post-event misinformation with a Self-Administered Interview. Appl Cogn Psychol, 26, 568-575.

Gabbert, F., Memon, A., Wright, D.B., 2006. Memory conformity: Disentangling the steps toward influence during a discussion. Psychon Bull Rev, 13, 480-485.

Garry, M., Manning, C.G., Loftus, E.F., Sherman, S.J., 1996. Imagination inflation: Imagining a childhood event inflates confidence that it occurred. Psychon Bull Rev, 3, 208-214.

Glanzer, M., Adams, K., 1985. The mirror effect in recognition memory. Mem Cognit, 13, 8-20.

Gordon, L.T., Shapiro, A.M., 2012. Priming correct information reduces the misinformation effect. Mem Cognit, 40, 717-726.

Greene, E., Flynn, M.S., Loftus, E.F., 1982. Inducing resistance to misleading information. Journal of Verbal Learning Verbal Behav, 21, 207-219.

Heaps, C., Nash, M., 1999. Individual differences in imagination inflation. Psychon Bull Rev, 6, 313-318.

Hekkanen, S.T., McEvoy, C., 2002. False memories and source-monitoring problems: Criterion differences. Appl Cogn Psychol, 16, 73-85.

Higham, P.A., 2002. Strong cues are not necessarily weak: Thomson and Tulving (1970) and the encoding specificity principle revisited. Mem Cognit, 30, 67-80.

Higham, P.A., Arnold, M.M., 2007. Beyond reliability and validity: The role of metacognition in psychological testing. In DeGregorio, R. A. (Ed.), New Developments in Psychological Testing (pp.139-162). New York: Nova Science Publishers, Inc.

Higham, P.A., Luna, K., Bloomfield, J., 2011. Trace-strength and source-monitoring accounts of accuracy and metacognitive resolution in the misinformation paradigm. Appl Cogn Psychol, 25, 324-335.

Higham, P.A., Perfect, T.J., Bruno, D., 2009. Investigating strength and frequency effects in recognition memory using Type-2 Signal Detection Theory. J Exp Psychol Learn Mem Cogn, 35, 57-80.

Holmes, A.E., Weaver, C.A., 2010. Eyewitness memory and misinformation: Are Remember/Know judgments more reliable than subjective confidence? Appl Psychol Crim Justice, 6, 47-61.

Johnson, M.K., Hashtroudi, S., Lindsay, D.S., 1993. Source monitoring. Psychol Bull, 114, 3-28.

Johnson, M.K., Raye, C.L., 1981. Reality monitoring. Psychol Rev, 88, 67-85.

Koriat, A., Goldsmith, M., 1996. Monitoring and control processes in the strategic regulation of memory accuracy. Psychol Rev, 103, 490-517.

Lindsay, D.S., 1990. Misleading suggestions can impair eyewitnesses’ ability to remember event details. J Exp Psychol Learn Mem Cogn, 16, 1077-1083.

Lindsay, D.S., Johnson, M.K., 1989. The eyewitness suggestibility effect and memory for source. Mem Cognit, 17, 349-358.

Loftus, E.F., 1979. Reactions to blatantly contradictory information. Mem Cognit, 7, 368-374.

Loftus, E.F., Miller, D.G., Burns, H.J., 1978. Semantic integration of verbal information into visual memory. J Exp Psychol Hum Learn, 4, 19-31.

McCloskey, M., Zaragoza, M., 1985. Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. J Exp Psychol Gen, 114, 1-16.

Meade, M. L., Roediger, H.L. III., 2002. Explorations in the social contagion of memory. Mem Cognit, 30, 995-1009.

Nelson, T.O., 1984. A comparison of current measures of the accuracy of feeling-of-knowing predictions. Psychol Bull, 95, 109-133.

Norman, D.A., Wickelgren, W.A., 1969. Strength theory of decision rules and latency in retrieval from short-term memory. J Math Psychol, 6, 192-208.

Perfect, T.J., Stollery, B., 1993. Memory and metamemory performance in older adults: One deficit or two? Q J Exp Psychol-A, 46, 119-135.

Perfect, T.J., Weber, N., 2012. How should witnesses regulate the accuracy of their identification decisions: One step forward, two steps back? J Exp Psychol Learn Mem Cogn, 38, 1810-1818.

Skagerberg, E.M., Wright, D.B., 2008a. The co-witness misinformation effect: Memory blends or memory compliance? Memory, 16, 436-442.

Skagerberg, E.M., Wright, D.B., 2008b. The prevalence of co-witnesses and co-witness discussions in real eyewitnesses. Psychol Crime Law, 14, 513-521.

Sondhi, V., Gupta, A., 2007. The quality of suggested memories. Int J Psychol, 42, 116-123.

Stretch, V., Wixted, J.T., 1998. On the difference between strength-based and frequency-based mirror effects in recognition memory. J Exp Psychol Learn Mem Cogn, 24, 1379-1396.

Thomas, A.K., Bulevich, J.B., Chan, J.C.K., 2010. Testing promotes eyewitness accuracy with a warning: Implications for retrieval enhanced suggestibility. J Mem Lang, 63, 149-157.

Thomas, A.K., Bulevich, J.B., Loftus, E.F., 2003. Exploring the role of repetition and sensory elaboration in the imagination inflation effect. Mem Cognit, 31, 630-640.

Tomes, J.L., Katz, A.N., 2000. Confidence-accuracy relations for real and suggested events. Memory, 8, 273-283.

Underwood, J., Pezdek, K., 1998. Memory suggestibility as an example of the sleeper effect. Psychon Bull Rev, 5, 449-453.

Wright, D.B., 1993. Misinformation and warnings in eyewitness testimony: A new testing procedure to differentiate explanations. Memory, 1, 153-166.

Wright, D.B., Self, G., Justice, C., 2000. Memory conformity: Exploring misinformation effects when presented by another person. Br J Psychol, 91, 189-202.

Zaragoza, M. S., & Koshmider, J. W. (1989). Misled subjects may know more than their performance implies. J Exp Psychol Learn Mem Cogn, 15, 246-255.

Zaragoza, M. S. & Lane, S. M. (1994). Source misattributions and the suggestibility of eyewitness memory. J Exp Psychol Learn Mem Cogn, 20, 934-945.

Footnotes

[1] We instructed our participants to write down a response for each question for three reasons. First, the calculation of Type-2 d' requires a response to be given for all test items (Higham, 2002). Second, all previous studies of resolution following misinformation exposure have used forced-report data to calculate the coefficients (Bonham and González-Vallejo, 2009; Cann and Katz, 2005; Higham et al., 2011; Tomes and Katz, 2000) thus allowing us to compare our results to those of prior studies directly. Third, the calculation of resolution requires at least one error is reported. If we had allowed participants to opt out of responding, we would have had insufficient data to calculate valid, reliable coefficients for more than a handful of our participants.

[2] Confidence ratings in Experiment 2 were recorded for exploratory purposes only, and are not reported here.

Table 1

Mean (and Standard Deviation) proportion correct, incorrect-guess, and incorrect-misinformation in Experiment 1

	Immediate		Delayed	
Response type	Control	Misleading	Control	Misleading
Correct	.78 (.10)	.67 (.17)	.65 (.11)	.55 (.13)
Incorrect-guess	.22 (.10)	.11 (.20)	.35 (.11)	.20 (.10)
Incorrect-misinformation	---	.22 (.16)	---	.26 (.13)

Table 2

Mean (and Standard Deviation) proportion of video, text, both, and neither responses to control and misleading items in Experiment 2

	Control Items		Misleading items	
Response	Immediate	Delayed	Immediate	Delayed
Video	.48 (.18)	.42 (.20)	.13 (.11)	.20 (.13)
Text	.03 (.07)	.09 (.08)	.34 (.17)	.15 (.12)*
Both	.23 (.22)	.28 (.16)	.21 (.13)	.30 (.15)
Neither	.26 (.15)	.21 (.16)	.33 (.16)	.35 (.15)
Memory parameter				
IM	.74 (.15)	.79 (.16)	.67 (.16)	.65 (.15)
SM	.67 (.26)	.52 (.22)	.50 (.22)	.24 (.20)

Note: * Difference between immediate and delayed conditions significant to Bonferroni-corrected $\alpha < .006$.

Figure Captions

Figure 1. Mean Goodman-Kruskal gamma coefficients in the immediate and delayed conditions in Experiment 1 Error bars represent 95% confidence intervals for within-subjects comparisons.

Figure 2. Mean Type-2 d' in the immediate and delayed conditions in Experiment 1. Error bars represent 95% confidence intervals for within-subjects comparisons.

Figure 3. Schematic representation of the process through which source monitoring influences item reporting and confidence judgments.

Figure 1.

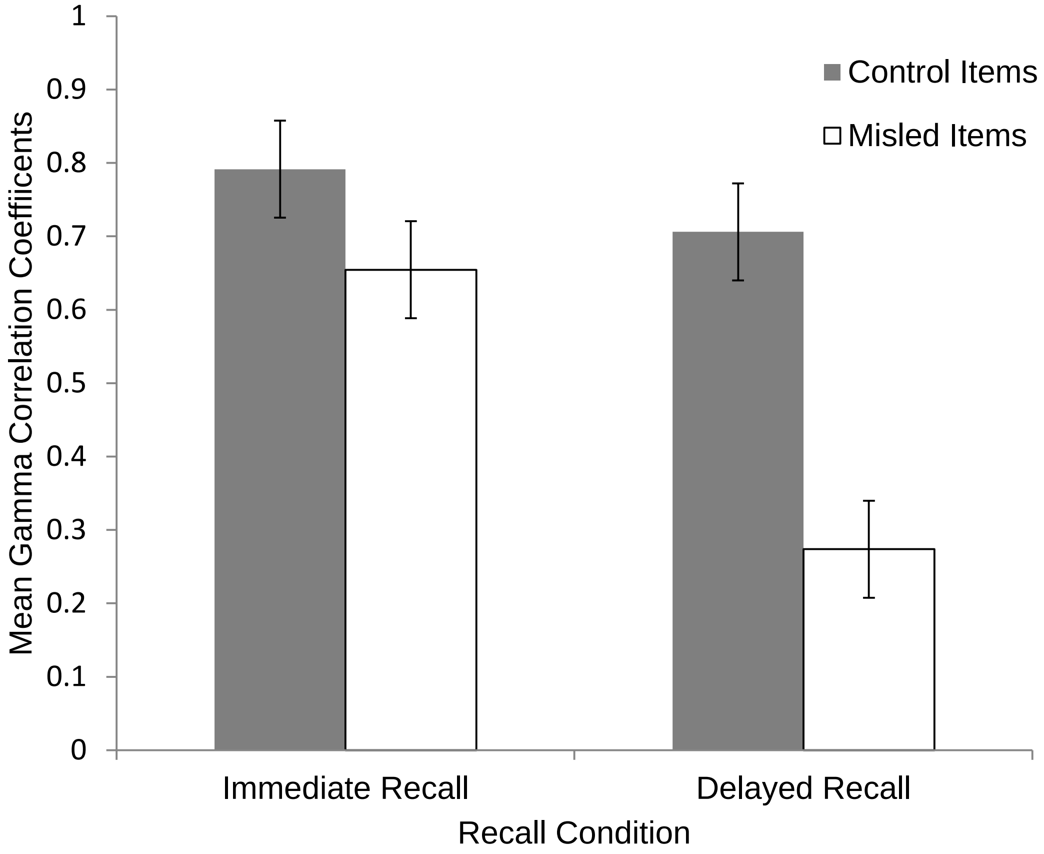

1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Mean Gamma Correlation Coeffiicents
Control Items
Misled Items
Immediate Recall
Delayed Recall
Recall Condition

Figure 2.

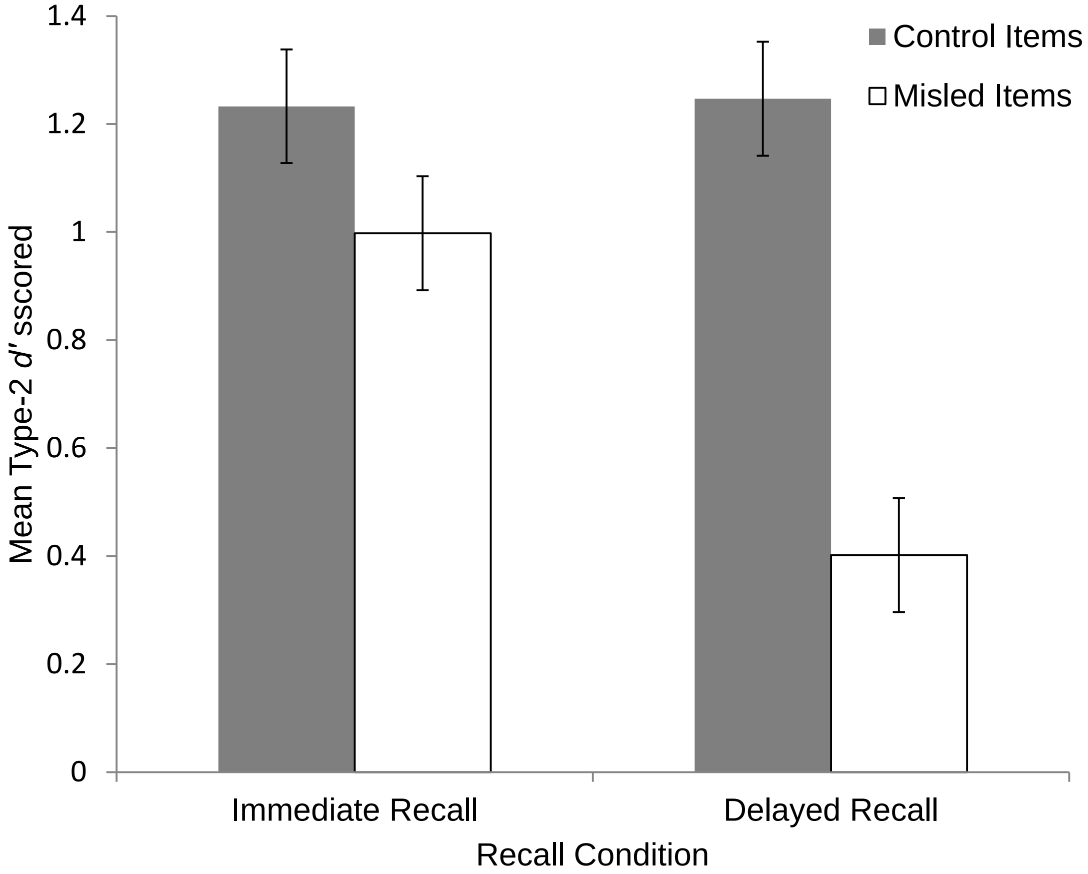
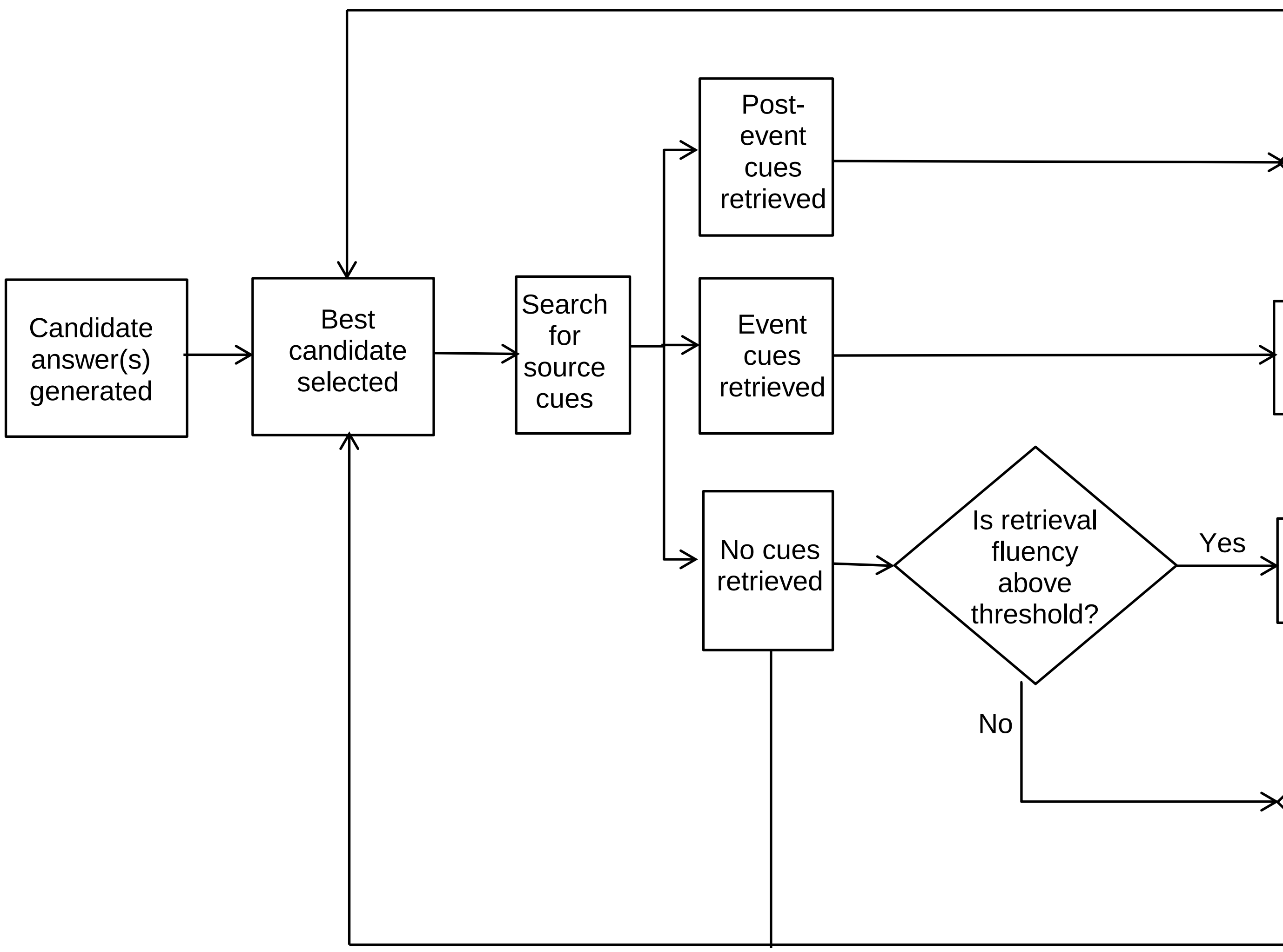

Control Items
Misled Items
Mean Type-2 d′ sscored
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Immediate Recall
Delayed Recall
Recall Condition

Figure 3.


Candidate answer(s) generated
Best candidate selected
Search for source cues
Post-event cues retrieved
Event cues retrieved
No cues retrieved
Is retrieval fluency above threshold?
Yes
No

Appendix

Narrative Versions A and B

Note: Misleading details are shown in **bold**. Consistent control items are shown in *italics.*

Narrative A:

An old man is sitting in his lounge room in a **leather chair**, reading a book. Directly across from him is a small cabinet, and next to this is *a television* in the left-hand corner. There is a large picture of a **green forest scene** hanging on the middle wall. The old man hears a knock at the door, ignores it, but then hears the doorbell and goes over to the front door. The old man has grey hair, and is wearing a jumper and **black pants.**

The old man opens the front door to find a young man, who is dressed in a **green jumper** with a light pink shirt underneath. He is Caucasian, and has **a goatee**. The young man first informs the old man that many old people die from **inhalation of wood fire smoke** every winter, and then goes on to offer *a free quotation on gas installation.* The old man seems confused, but then lets him in.

The young man makes his way over to the red-brick fireplace and sets down his toolbox, which has a **red handle and latch**. He takes a **tape-measure** out from his tool box, and proceeds to inspect the fireplace. **He then asks the old man if he has the warranty papers** for the chimney chase**.** The old man says he shall have to go and search for it. While the old man is gone, the young man steals war memorabilia from the mantelpiece.

The old man is in his study, which has **green carpet and curtains,** searching for the papers. Meanwhile, in the lounge room, the young man sees the old man's watch sitting on the table next to the old man's book. He picks up the watch, looks at it, **and then just puts it back down on the table.** The old man then brings the papers out to the young man, who thanks him, and then asks for a drink. The old man says of course, and then goes to the kitchen.

The young man then searches the cupboard next to the fireplace, and inside are a lot of board-games and books. He first takes out a small grey box, and opens it to find **a silver ring inside.** He puts it in his toolbox, and continues to search the cupboard. He then finds a **silver jewellery box**, and from it he takes *various kinds of bracelets* and puts them in his toolbox.

The old man is in the kitchen, preparing the drinks. The **two yellow mugs** and plate are on a serving tray. Meanwhile, in the lounge room, the young man **finds a stash of coins** in *a money tin.* When the old man returns to the lounge room, the young man says that he is sorry, **just realized that he's late for an appointment** and has to go, and then quickly leaves the house.

Narrative B:

An old man is sitting in his lounge room, reading a book. Directly across from him is a small cabinet, and next to this is **a bookcase in the left-hand corner**. There is a large picture hanging on the middle wall. The old man hears a knock at the door, and ignores it. He then hears the doorbell, and makes his way to the door. The old man has grey hair and is **wearing a blue jumper** with a shirt underneath.

The old man **unlocks the gold safety chain,** and opens the front door to find the young man, who is dressed in a jumper, **denim jeans** and work boots**.** He is Caucasian, and has **messy brown hair.** The young man first informs the old man that *30,000 old people die from the cold every winter*, and then goes on to **offer a free fireplace service**. The old man seems confused, but then lets him in.

The young man makes his way over to the red-brick fireplace, and sets down his **red toolbox,** and proceeds to inspect the fireplace. Inside the fireplace is a **pile of firewood** sitting in the cast-iron grate. He then asks the old man if he has the *receipt for the chimney* *chase*. The old man says he shall have to go and search for it. While the old man is gone, the young man steals **an antique clock and a gold photo frame** from the mantelpiece.

The old man is in his study, searching for the papers. Meanwhile, in the lounge room, the young man sees the old man's watch sitting on the table next to the old man's **brown book.** He picks it up, and then quickly *puts it in his toolbox*. Back in the study, the old man finds the receipt **in a tray of papers on top of the desk**. He takes it out to the young man, who thanks him, and then asks for a drink. The old man says of course, and **offers him a coffee**.

The young man then searches the cupboard next to the fireplace. The cupboard doors are **cream coloured,** and inside are a lot of board-games and books. He first takes out a small grey box, and opens it. He puts it in his toolbox, and continues to search the cupboard. He then finds a jewellery box, and from it, **he takes a handful of gold and silver rings** and puts them into his toolbox.

The old man is in the kitchen, preparing the drinks and also plate of **assorted cakes.** Meanwhile, back in the lounge room, the young man finds a stash of money **in a brown envelope** in the cupboard. When the old man returns to the lounge, the young man says that he has an *emergency call out and has to go*, and then quickly leaves the house.