Running head: DISCUSSION AND EYEWITNESS MEMORY

Can a witness report hearsay evidence unintentionally? The effects of discussion on

eyewitness memory

Helen M. Paterson (The University of Sydney)

Richard I. Kemp (University of New South Wales)

Sarah McIntyre (The University of Sydney)

Address correspondence to:
Helen M. Paterson
School of Psychology
The University of Sydney
Sydney, 2006
Australia
E-mail: helen.paterson@sydney.edu.au
Phone: (02) 9036 9403
Fax: (02) 9036 5223

Abstract

When eyewitnesses are exposed to misinformation about an event from a co-witness, they often incorporate this misinformation in their recall of the event. The current research aimed to investigate whether this memory conformity phenomenon is due to change in the witness's memory for the event, or to social pressures to conform to the co-witness's account. Participants were shown a crime video and then asked to discuss the video in groups, with some receiving misinformation about the event from their discussion partners. After a one week delay some participants were warned about possible misinformation before all participants provided their own account of the event. In Study 1, participants made remember/know judgments about the items recalled, and in Study 2 they indicated the source of their memories. Co-witness information was incorporated into participants' testimonies, and this effect was not reduced by warnings or source monitoring instructions, suggesting memory change may have occurred. However, there was some indication that remember/know judgments may help distinguish between 'real' memories and co-witness information.

Can a witness report hearsay evidence unintentionally? The effects of discussion on eyewitness memory

Within the legal system it is commonly held that the most reliable eyewitness evidence is that given by independent eyewitnesses who have not communicated with one another (e.g., Heaton-Armstong, 1987). Consequently, many legal procedures are designed to prevent eyewitnesses from discussing the crime with one another. For example, a survey of police officers has revealed that they often attempt to separate witnesses and discourage them from talking about the event with one another (Paterson & Kemp, 2005). Despite these attempts, it is clear that witnesses often do talk to each other about the event they saw. Recent studies have found that the majority of eyewitnesses reported discussing details of the event with their co-witness (Paterson & Kemp, 2006a; Skagerberg & Wright, 2008).

Discussion between witnesses is problematic because research has shown that exposure to co-witness information about an event after it has occurred often causes people to incorporate this information into their accounts of the event. In fact, research has shown that co-witness information presented in this way is a more influential method of presenting misinformation than leading questions or written postevent narratives (Paterson & Kemp, 2006b). Similarly, another study has shown that postevent information encountered through co-witness discussion was significantly more influential than that encountered through a non-social source (Gabbert, Memon, Allan, & Wright, 2004). This phenomenon has become known as 'memory conformity' (e.g., Gabbert, Memon, & Allan, 2003; Wright, Self, & Justice, 2000). Research has shown that participants who have discussed an event with a co-witness who supplies misinformation show less accurate memory for the misled items than non-misled items (e.g., Hoffman, Granhag, See & Loftus, 2001; Schneider &

Watkins, 1996; Shaw, Garven, & Wood, 1997) and their memory is also less accurate for misled items than individuals who did not discuss the event (e.g., Gabbert, Memon, Allan, & Wright, 2004; Hollin & Clifford, 1983). Despite this clear evidence of co-witness memory conformity, we do not know why witnesses often report information provided by a co-witness. Some researchers have attempted to address this issue (e.g., Azad, Lindsay& Brimacombe, 2010; Gabbert et al., 2007; Paterson & Kemp, 2006), however it remains unclear whether the conformity is due to memory distortion or other factors.

Memory conformity could occur without any memory distortion if the participant reports co-witness information for other, more social, reasons. For example, an individual may conform in order to gain social approval (normative social influence), or they may report the second-hand information because they believe it to be correct (informational influence; Deutsch & Gerard, 1955). These mechanisms require that witnesses remember the co-witness information and that it was obtained from a second-hand source. For this reason, Betz, Skowronski, and Ostrom (1996) refer to these mechanisms as "source-tagged mechanisms." An alternative explanation for memory conformity suggests that memory change does occur. For example, according to the source monitoring theory (e.g., Johnson, Hashtroudi, & Lindsay, 1993; Lindsay & Johnson, 1989), a witness may remember the information provided by the co-witness, but forget that it was obtained second-hand. Thus, source-monitoring theory suggests that people report co-witness misinformation because they misremember the source of information they recall.

The investigation as to why participants often report co-witness information has important legal implications. If witnesses are unable to distinguish what they actually experienced from information obtained from a co-witness, then they may, in

effect, be reporting hearsay evidence without knowing it. The legal concept of 'hearsay' embodies the notion that "a witness's assertions of relevant facts should be based upon his or her own experiences" (Forbes, 2003, p. 59) and not those of another. However, implicit in this concept is the assumption that witnesses are able to distinguish their own experiences from information they obtain second-hand (e.g., from a co-witness). That is, within the legal system there is an assumption that witnesses would only report hearsay evidence because of social influence and not because of memory change, but this may not be the case. If witnesses are unable to distinguish their 'real' memories from second-hand information then their testimonies are considered 'contaminated' and this may lead to the unintentional presentation of hearsay evidence.

Researchers have employed three techniques in an attempt to identify the mechanism responsible for memory conformity. The first method has been to ask participants to make a remember/know judgment (Tulving, 1985) for each item they recall following the discussion. Roediger, Meade and Bergman (2001) had participants recall items from images of common household scenes in alternation with a confederate who sometimes recalled items that were not in the scenes. Participants were asked to produce a remember/know judgment for each item they recalled. If they consciously remembered seeing the object in the scene they would indicate that they "remembered" it. In contrast, "know" responses were used to indicate items for which the participants did not have any specific recollection, but which they believed to be in the scene. When participants incorrectly reported items mentioned by the co-witness, they were more likely to claim that they "knew" the suggested items had been in the scenes than to report that they specifically "remembered" seeing them there. These results were replicated in studies by Meade and Roediger (2002). This

implies that participants are at least partially able to discriminate between information obtained from a co-witness and that obtained first-hand from the scene.

The second experimental technique used to determine the mechanism most likely responsible for memory conformity is to warn participants that they may have been exposed to misinformation from their co-witness. Meade and Roediger (2002) found that the warnings significantly reduced the effect of the co-witness misinformation, but did not eliminate it, suggesting that some distortion of the original memory occurred. In another study, Wright et al. (2008) found that strict warnings (i.e., participants were told to recall items only if they were sure that they were accurate) reduced the effects of co-witness misinformation, but also had the deleterious effect of reducing the number of accurate details recalled.

The third technique employed to identify the mechanism responsible for memory conformity is source monitoring (Lindsay & Johnston, 1989). Source monitoring tests give participants the option of saying that they remember an item only from the postevent suggestion, and not from the original stimulus. For example, participants in the current study were asked to indicate whether they remembered the item from the video only, from the discussion only, or from both the video *and* discussion. Although some studies have shown that source monitoring instructions such as these can eliminate the standard misinformation effect (Lindsay & Johnson, 1989; Zaragoza & Koshmider, 1989), other evidence suggests that co-witness contamination persists even when participants are asked to monitor the source of the information they recall (Meade & Roediger, 2002). In fact, Gabbert, Memon and Wright (2007) found that participants errantly attributed the source of their memory approximately 50% of the time.

While the studies described above (Meade & Roediger, 2002; Roediger et al., 2001) provide valuable insights into the mechanisms underlying memory conformity, they do have some limitations. First, in these studies there was only a short time delay (i.e. within the same testing session) between the presentation of misinformation and the memory testing. Real witnesses are commonly interviewed after longer delays (Wright & McDaid, 1996). This is important because source misattribution is more likely to occur after a longer delay. Underwood and Pezdek (1998) found that source information was less accurately reported after a delay of one month than after a ten-minute delay. Thus, warnings and source monitoring instructions might be less effective if the duration of the delay more accurately reflected real eyewitness situations. Another limitation of these studies is their use of static photos as stimuli, which may reduce the ecological validity. Furthermore, in the studies described above (Meade & Roediger, 2002; Roediger et al., 2001), participants took turns recalling items, which restricted their communication in an artificial way. In particular, the social influence and persuasion likely to be part of real co-witness communication was absent. Finally, the ecological validity of co-witness studies is compromised by the use of confederates to supply the false information. Although this method allows the experimenters control over the false information which the participant is exposed to, confederates may not act like genuine co-witnesses. For example, confederates may have total confidence in the misinformation they supply, and may take a more dominant social role than real co-witnesses would.

The present research aimed to improve on these limitations and increase ecological validity by having a longer delay period between co-witness discussion and individual testing, using videos to present a more realistic crime scenario, and by allowing more natural conversations between pairs of genuine witnesses. We

achieved this by showing pairs of participants what they believed were identical videos, but which actually differed in some details. Participants were then asked to discuss the stimulus event with one another in groups, some of which contained members who had seen different versions of the stimuli. Because the participants were unaware that they had seen different stimuli, their interactions with the group were likely to be more natural than those of delegated confederates. This methodology has been used in several studies of the effects of co-witness discussion on memory (e.g., Gabbert, Memon, & Allan, 2003; Garry, French, Kinzett, & Mori, 2008; Paterson, Kemp, & Ng, 2009; Wright, Self, & Justice, 2000). These studies have largely confirmed the results of studies using confederates, finding that participants who viewed different versions of the video erroneously recalled items from the alternative version of the stimulus when later giving their individual accounts.

However, many studies using this improved methodology have not employed appropriate control groups. For example, some studies have compared participants who have seen different versions of the stimulus with individuals who have not discussed the event. This approach confounds the effects of discussing the event (e.g., the social presence of another, the elaboration and rehearsal of event-related memories) with the influence of the co-witness (i.e., the introduction of misinformation). There are also confounds when all participants have seen different versions of the stimulus than their partners and a comparison is made between items that differed between the two stimuli, and those that did not (e.g., carryover effects associated with within-subjects design). To improve on this methodology, we used "same-video" control groups (or "natural discussion" control groups as we will refer to them) in which all members of the group had seen the same version of the video

(either all Version A or all Version B). Using natural discussion groups as a control is critical because it allows us to distinguish between the effects of discussion and the effects of co-witness misinformation. Furthermore, it also allows us to identify positive as well as negative effects of discussion. For example, research on collaborative inhibition has shown that collaborative discussion groups typically recall less than non-interacting groups of pooled individuals (Weldon, Blair, & Huebsch. 2000). However, such group discussion can benefit memory when the group members are later tested individually. That is, people who had been collaborating in groups performed better on a final individual free recall task than those who had been in non-interacting nominal groups (Basden, Basden, & Henry, 2000).

These natural discussion groups that do not involve the experimental introduction of postevent information (either by a confederate or any other method), have potentially the highest ecological validity, and yet few applied studies have employed such controls. Furthermore, those studies which have employed this natural discussion control group have yielded inconsistent results. Results from some experiments which do not artificially introduce postevent misinformation suggest that under certain conditions, group recall can have a beneficial effect on eyewitness testimony (e.g., Underwood & Milton, 1993; Yarmey & Morris, 1998), while others suggest that discussion prior to individual recall is not an advantageous procedure (e.g., Paterson, Kemp, & Ng, 2009; Stephenson, Abrams, Wagner, & Wade, 1986; Yarmey, 1992).

The present research aimed to investigate effects of warnings, remember/know judgments and source monitoring on co-witness memory conformity following a one-week retention interval and using an ecologically valid design that incorporates a Natural Discussion control group. Experiment 1 investigated the effects of warnings

and remember/know judgments on memory conformity and Experiment 2 investigated the effects of warnings and source monitoring judgments on memory conformity.

Experiment 1

The first objective of Experiment 1 was to determine whether memory conformity is due to memory distortion or to other 'source-tagged mechanisms' (Betz et al., 1996). To achieve this, participants in Experiment 1 received warnings that they may have been exposed to misinformation. These explicit warnings about misinformation were designed to help to reduce the experimental demand that may encourage conformity. It was hypothesised that, consistent with other studies (Betz et al., 1996; Gallo, Roberts & Seamon, 1997; Gallo, Roediger & McDermott, 2001; McDermott & Roediger, 1998; Meade & Roediger, 2002; Wright, 1993), memory conformity would be reduced, but not eliminated by the warning. Participants were also asked to make a remember/know judgment (Tulving, 1985) for each detail reported. In accordance with previous research (Meade & Roediger, 2002; Roediger et al., 2001) it was hypothesized that participants would be more likely to claim that they "knew" the items mentioned by the co-witness than to report that they specifically "remembered" seeing them.

The second objective of the current study was to compare the memory of individuals who discussed the event with a co-witness who had seen a slightly different version of the eyewitness stimulus with the memory of individuals in natural discussion groups (who all saw the same stimulus), and also with individuals who had not discussed the event. In accordance with previously demonstrated results (e.g., Gabbert et al., 2001; Garry et al., 2008; Wright et al., 2000), it was hypothesized that individuals who discussed the event with a co-witness who had seen a different

version of the video would show less accurate memory than those in the natural

discussion groups or those who did not discuss the event.

Method

Participants

Participants were 119 undergraduate psychology students (97 female, 22

male), with an average age of 20.9 (*SD* = 1.40) years. Participants were invited to take

part in the experiment during a class tutorial. Informed consent was obtained and they

received no incentives for their participation in the study.

Design

After viewing a crime video, participants were randomly allocated to one of

two Discussion Type conditions (discussion with co-witnesses who all saw the same

version of the video, discussion with co-witnesses who saw different versions) and

one of three Warning conditions (no warning about misinformation, specific warning,

general warning). In addition there was a 'No Discussion' control group. Thus, the

study employed a 2 x 3 + 1 between subjects design. The dependent variables were

memory accuracy (measured using recognition questionnaires and free recall) and

remember/know judgments.

Materials

Eyewitness stimuli. The eyewitness stimulus used in this study was a short

video (under two minutes in duration) depicting a robbery. There were two slightly

different versions of the video (Version A and Version B). Table 1 summarises the

differences between the videos. The videos were displayed to the participants on

individual computer monitors. Each participant viewed only one of the two videos,

but participants were led to believe they were all viewing the same video.

Free recall. Participants were given as much time as they needed to recall everything they could remember about the crime video. Specifically, they were asked to write about the sequence of events, what was said, the setting, what the people looked like, what was stolen from the flat, and any other details they could remember. This questionnaire was filled out by participants in the control condition in the first session (while participants in the discussion conditions talked about the incident) and all participants in the second session.

Recognition questionnaire. A recognition questionnaire containing 23 true/false questions about the video was developed. Of the 23 questions, 18 concerned *non-misled items* (i.e., none of the participants received misleading information regarding these items during the group discussion). The remaining five questions concerned *misled items* (i.e., false statements for which some people received misinformation from their co-witness). These five items regarded aspects which differed between Version A and Version B of the video. For each question, participants were asked to indicate whether the statement was true or false and then for each "true" response to make a "remember" or "know" judgment (Tulving, 1985). Participants were given definitions of "remember" and "know" based on instructions given by Rajaram (1993). That is, participants were instructed, "If your recognition of the item is accompanied by a conscious recollection of its occurrence in the video, then state that you "Remember" it. "Remember" is the ability to become consciously aware again of some aspect or aspects of what happened or what was experienced at the time the video was presented (e.g., aspects of the physical appearance of the item in the video, or of what you were thinking at the time). In other words, the "remembered" statement should bring back to mind a particular association, image, or something more personal from the time of viewing the video, or something about its

appearance or position. "Know" responses should be made when you believe that the item was in the video, but you cannot consciously recollect anything about its actual occurrence or what happened or what was experienced at the time of its occurrence. In other words, indicate, "know" when you believe the statement to be true, but it fails to evoke any specific recollection from the video."

Warnings. One third of the participants were given a 'specific warning' that some students in their group may have seen a slightly different video and that they should answer the questions solely on the basis of what *they* themselves remember from the video. This warning was designed to decrease the credibility of the co-witness, so that participants would attempt to rely on their own memories for the video. Another third of the participants were given a more 'general warning' that they should make a specific effort to disregard what others in their discussion group told them and answer the questions solely on the basis of what *they* themselves remember from the video. This warning was designed to loosely reflect a hearsay warning that witnesses may receive. The remaining participants received no warning.

Manipulation awareness check. Participants in discussion groups were asked four questions relating to their discussion experiences in order to determine whether those in the different-video group had similar experiences to those in the same-video groups (i.e., how useful they felt the discussion was in helping them remember the video clip, whether they felt they had learned any new information during the discussion, how freely they believe their discussion group had exchanged ideas, and whether they think their answers were influenced by the group discussion). Participants were asked to indicate their response to each question using a six-point scale.

Procedure

This study took place during two sessions, one week apart. During the first session the students gathered in a computer laboratory and were each seated at a computer to view the crime video. After informed consent was obtained, participants were instructed that they would each be viewing the same crime scene on the individual computer terminals. Unbeknownst to the participants, each computer had been loaded with one of the two different versions of the video.

After viewing the video, there was a delay of twenty minutes (during which the participants did their regular class work). Some of the participants were then split into groups of four to discuss the crime video. The groups were arranged so that half contained individuals who had all seen the same version of the video (same-video groups) while the other groups had an even mix of participants who had seen Version A and Version B (different-video groups).

Participants in the control group were asked not to discuss the video with anyone else, but instead were asked to write about the video individually. Control participants were in a separate room so that they could not overhear any of the discussions taking place. All participants were given ten minutes to discuss or write about the video using a guided recall procedure. Specifically, the participants were instructed to discuss/write about the sequence of events, the setting, what the people looked like, what was stolen from the flat and any other information they could remember. At the end of ten minutes, participants were asked not to discuss the event any further.

One week later, participants returned to the lab for the second session. Participants in the discussion groups (both same- and different-video groups) were asked to answer the manipulation awareness questions about their experiences of the

group discussion. All participants were then asked to state what they believed to be the purpose of the experiment.

Participants in the discussion conditions were then randomly assigned to one of three warning conditions. One third of participants were given the specific warning, another third were given the general warning, and the remaining third were given no warning about potential misinformation. After the warning manipulation, all participants were individually tested for their recall of the event. First, they were given the free recall task where they wrote down as much information as they could remember about the video. They were then asked to complete the true/false recognition questionnaire with remember/know judgments. Finally, participants were fully debriefed, thanked for their participation, and asked not to discuss the experiment with other potential participants.

Results and Discussion

Manipulation Awareness Check

The participants' responses to questions about their group discussion experiences were analysed. Four independent t-tests revealed that there were no significant differences between the two discussion groups for any of the four discussion questions, all p's > .05. This implies that participants in different-video groups were unaware of the manipulation, as they reported their discussions to be just as "influential, informative, open, and free" as the participants in same-video groups.

When asked to state the purpose of the study, only one participant gave an answer which suggested that he thought he might have been shown a slightly different video from the other members of his group. However, this particular participant was in a same-video group, so the experimental manipulation was considered successful and no participants were excluded from the analyses.

Memory Accuracy

Recognition questionnaire. From the recognition questionnaire, total accuracy scores were calculated for the eighteen items that were not subject to misinformation (i.e., non-misled items) and a separate score was calculated for the five items for which we had attempted to mislead some participants (i.e., misled items). The mean scores for memory accuracy are shown in Figure 1.

Separate ANOVAs were conducted to determine the effect of Discussion Type (different-video, same-video, no discussion) on memory accuracy for *misled* and *non-misled items*. There was a significant main effect of Discussion Type on *misled items*, F (2, 114) = 5.04, $p < .01$; partial eta squared = .08). Bonferroni post hoc tests revealed that participants in the different-video group were less accurate on misled items ($M = 58.95$, $SD = 21.85$) than either the participants in same-video groups ($M = 70.67$, $SD = 17.99$; $p < .05$) or the no discussion controls ($M = 72.00$, $SD = 22.65$; $p < .05$). There was no difference between the same-video group and no discussion controls on accuracy for misled items ($p = 1.0$). Analyses of the individual misled items also revealed that participants in the different-video group had a consistently high error rate compared to the others groups, however on its own only one item (thief was drinking beer) reached significance, χ^2 (2, $N = 118$) = 25.87, $p < .001$. These results support the hypothesis that participants who discussed the video with people who had seen an alternate version erroneously incorporated information from the alternate version into their own individual account. That is, they demonstrate the presence of a co-witness memory conformity effect.

Contrary to our expectations, there was also a main effect of Discussion Type on *non-misled items*, F (2, 115) = 5.13, $p < .01$; partial eta squared = .08). Bonferroni post hoc tests revealed that participants in the different-video group were significantly

more accurate on non-misled items ($M = 81.42$, $SD = 9.40$) than either those in the same-video group ($M = 75.56$, $SD = 9.85$, $p < .05$) or the no discussion controls ($M = 75.37$, $SD = 11.55$, $p < .05$). The difference between the same-video group and no discussion controls on accuracy for non-misled items was not significant ($p = 1.0$). Possible explanations for this finding are examined in more detail in the General Discussion section.

Separate 2 x 3 ANOVAs were conducted to investigate the effects of Discussion Type and Warning on participants' accuracy for *non-misled items* and *misled items* (this analysis excluded the no-discussion control group because it did not make sense to give this group a warning when they had not discussed the video with anyone). There was no main effect of Warning on *misled items*, F (2, 81) = 0.72, $p = .49$) and no interaction between Discussion Type and Warning, (F (2, 81) = 0.42, $p = .69$). There was also no main effect of Warning on *non-misled items*, F (2, 82) = 0.88, $p = .42$) and no interaction between Discussion Type and Warning, (F (2, 82) = 0.79, $p = .45$).

The fact that the warnings did not affect memory conformity could suggest that participants were unable to distinguish what they actually witnessed from the information they discussed with the co-witness. Alternately, it could be the case that the general and specific warning conditions in this study were too similar to each other in that they both instructed the participant to disregard the PEI from their partner, and answer the questions based on their own memory only. However, when we tried combining these two warning conditions and comparing them with the no warning control group we still found no effect of Warning and no interaction between Warning and Discussion Type (all p's >.25). In Experiment 2 we eliminated the the

general warning condition, giving participants either a specific warning or no warning.

Free recall. Four judges naïve to the experimental conditions independently coded the free recall narratives. Each narrative was scored in terms of: (a) accurate information (regarding sequence of events, character descriptions, dialog, setting, and items stolen) and (b) misleading postevent information reported. The participants' responses were randomly assigned to one of the four coders. Twenty of the free recall narratives were scored by all coders in order to assess inter-rater reliability. There was a significant correlation between the four coders on *accurate information* and *misleading postevent information* (*r*'s ranged from .80 to .95, all *p*'s <.01). The mean scores for memory accuracy are shown in Figure 2.

Separate ANOVAs were conducted to determine the effect of Discussion Type (different-video, same-video, no discussion) on *accurate information* and *misleading postevent information.* There was a main effect of Discussion Type on *misleading postevent information*, $F(2, 115) = 5.04$, $p < .01$; partial eta squared = .08. Bonferroni post hoc tests revealed that participants in the different-video group were more likely to report misleading postevent information ($M = 0.32$, $SD = 0.68$) than participants in same-video groups ($M = 0.00$, $SD = 0.00$, $p < .05$). There was also a significant main effect of Discussion Type on *accurate information*, $F(2, 115) = 7.02$, $p < .01$; partial eta squared = .11. Bonferroni post hoc tests revealed that participants in the different-video group were significantly more likely to report accurate propositions ($M = 49.81$, $SD = 13.19$) than participants in the control condition ($M = 40.72$, $SD = 10.76$, $p < .05$) and the same-video group ($M = 43.63$, $SD = 7.65$, $p = .05$).

Separate 2 x 3 ANOVAs were conducted to investigate the effects of Discussion Type and Warning on *accurate information* and *misleading postevent*

information (this analysis excluded the no-discussion control group). There was no main effect of Warning on recalled *misleading postevent information*, F (2, 83) = 0.33, p = .72 and no interaction between Discussion Type and Warning, F (2, 83) = 0.33, p = .72. There was also no main effect of Warning on recalled *accurate information, F* (2, 83) = 1.31, p = .27 and no interaction between Discussion Type and Warning, F (2, 83) = 2.04, p = .14).

In sum, analysis of free recall accuracy showed the same pattern of results as the recognition questionnaire for both the Discussion Type and Warning manipulations. That is, participants in the different-video group were more likely to recall misleading postevent information as well as accurate information, and warnings did not have any effect on memory conformity.

Percentage of Participants Who Reported Misinformation on Free Recall. The accuracy scores were recoded to indicate those participants who, during free recall, mentioned at least one item of misleading postevent information. It was found that 22% of participants in the different-video group reported at least one misled item, while only 10% in the control group and 0% in the same-video group inaccurately reported at least one misled item. A chi-square analysis revealed a significant association between Discussion Type and whether or not participants mentioned at least one item of misleading postevent information (χ^2 (2, N = 119) = 8.71, p < .05).

Remember/Know Judgments

In our first analysis of Remember/Know (R/K) judgments, only data from those in the different-video group were included. We analysed the items that half of the participants had been exposed to in their version of the video (i.e., those who had seen Version A of the video), and the remaining half (those who saw Version B) had not been exposed to in their version of the video (and hence about which they were

likely to have heard misinformation during their group discussion). We refer to these two groups as the non-misled and misled groups respectively. R/K responses were only given if the participants responded that the statement was "true", and thus for these items participants in the non-misled group provided R/K judgements in relation to correct responses regarding items about which they were not subject to misinformation, while misled participants provided R/K judgements to erroneous responses to items which were subject to misinformation.

Non-misled participants indicated "remember" an average of 76.9% of times ($SD =$22.54) when correctly responding to the items, while the misled participants made "remember" responses only 41.6% of occasions ($SD = 34.08$) when erroneously responding to the items, a statistically significant difference (t (56) = 4.73, $p < .0005$; see Figure 3). The 95% confidence intervals of the difference suggest that non-misled participants were between 20.4% and 50.3% more likely to report they "remembered" a detail than were the misled participants. A similar pattern emerged when this analysis was repeated at the item level. That is, for each item non-misled participants were more likely to make a "remember" response (and hence less likely to make "know" responses) than were misled participants.

In a separate analysis that included all participants, we calculated the proportion of times that participants were correct when they responded "remember" and the proportion of times they were correct when they reported "know" (see Figure 4). A paired samples t-test showed participants were significantly more likely to be accurate if they indicated "remember" ($M = 90.82$, $SD = 11.77$) than if they indicated "know" ($M = 72.53$, $SD = 23.13$; t (116) = 8.42, $p < .0005$; partial eta squared = .379). Inspection of the confidence intervals indicated that participants were between 14 and

23% more likely to be accurate if they indicated 'remember' than if they indicated 'know.'

These results indicate the Remember/Know Judgements may be of some value as an indicator of accuracy. This is consistent with the findings of Roediger et al. (2001) who found that when participants reported misinformation, they were more likely to report they "knew" the misled item was in the original scene, rather than reporting they consciously "remembered" it was there. Roediger et al. use source monitoring theory (e.g., Johnson et al., 1993; Lindsay & Johnson, 1989) to explain this finding. They suggest that most reported misinformation tends to be "known" rather than "remembered" because participants have difficulty locating the source of their memories. They argue that if a source could be confidently attributed, the item should be judged as "remembered." The results from the remember/know judgments suggest that it may be possible to distinguish between 'real' memory and information obtained from a co-witness by tapping in to witnesses' source attribution.

Although overall in Experiment 1, people were more likely to report that misinformation was "known" rather than "remembered" the majority of participants (67%) reported that they consciously "remembered" at least one item of misinformation. Other studies have also shown that misinformation can be accompanied by remember judgments (e.g., Higham & Vokey, 2004; Roediger & McDermott, 1995). When false memories are accompanied by "remember" judgments, this suggests a memory change may have occurred (Meade & Roediger, 2002), limiting the predictive power of remember/know judgements for memory accuracy.

In this study, we have extended our understanding of remember/know judgments by demonstrating in a more ecologically valid testing environment that

they manifest in a way that might be predicted from previous research. We have also shown that the use of a Natural Discussion control group yields similar results to the use of conventional no discussion control groups, but that the Natural Discussion control is to be preferred because of greater ecological validity and the ability to reveal the positive effects of discussion in the absence of introduced misinformation, as was observed in Experiment 1.

Experiment 2

Results of Experiment 1 suggest that it may be possible to use remember/know judgments (Tulving, 1985) to distinguish between real memories and those obtained second-hand. While this finding has important theoretical implications, it has limited practical implications given its modest predictive power. The practicality of using this judgment as a way of distinguishing real from confounded memory may also be diminished by the difficulty in explaining the distinction between a "remember" and a "know" state to eyewitnesses.

In Experiment 2, we investigated whether a procedure that was easier to explain might more accurately distinguish between real memories and those obtained from a co-witness. Participants were given a source monitoring task in which they were asked to report the context in which they remembered encountering each "true" item: when watching the video of the event, during the discussion with co-witnesses, or both. This procedure was based on source monitoring instructions developed by Lindsay and Johnson (1989).

It was unclear whether the source monitoring instructions would reduce memory conformity, since it has not yet been tested in an ecologically valid environment using crime videos as stimuli and genuine co-witnesses. Lindsay and

Johnson (1989) and Zaragoza and Koshmider (1989) both found it eliminated the standard misinformation effect, whereas Meade and Roediger (2002) found the effect persisted even when participants were asked to monitor the source of the information they recalled.

In Experiment 2 we simplified our design by eliminating the 'No Discussion' control group because it was not directly relevant to the hypotheses.   In the introduction we outlined a number of reasons why the Natural Discussion Group provides a more appropriate and ecologically valid control group for these studies. The elimination of the No Discussion condition also meant we could use a full factorial design. Experiment 2 was further simplified by eliminating the general warning condition, and giving participants either a specific warning or no warning.

Because memory conformity may be reduced when participants are given a warning (e.g., Betz, Skowronski, & Ostrom, 1996; Gallo, Roberts & Seamon, 1997; Gallo, Roediger & McDermott, 2001; McDermott & Roediger, 1998; Wright, 1993) or when they are given a source monitoring test (Lindsay & Johnson, 1989) it is possible that even if neither of these potential mitigating factors eliminates the conformity effect alone, both together may do so.

Method

Participants

Participants were 64 undergraduate psychology students (48 female, 16 male), with an average age of 20.9 ($SD = 2.32$) years. Participants were invited to take part in the experiment during a class tutorial. They received no incentives for their participation in the study.

Design

This experiment employed a 2x2 full factorial between-subjects experimental design, which investigated the effects of Discussion Type (same-video group, different-video group) and Warning (no warning, specific warning) on memory for misled and non-misled items. The dependent variables were memory accuracy and participants' judgments about the source of their memory.

Materials

Eyewitness stimuli. Experiment 2 used the same materials used in Experiment 1.

Recognition questionnaire. Experiment 2 employed the same 23 questions as Experiment 1, however the instructions were changed slightly. As in Experiment 1, participants were asked to indicate whether the statement was true or false. However, instead of making remember/know judgments, participants were asked to make a source monitoring judgment. That is, for the items they said were "true", they were asked to indicate whether they were remembered from: (1) the video only, (2) the discussion only, or (3) both the video and the discussion.

Procedure

Experiment 2 followed the same procedure as Experiment 1 except for the following minor differences. Half of the participants in each level of the Discussion Type factor were randomly allocated to the 'specific warning' condition, and the other half were allocated to the 'no warning' condition. Once participants had completed the memory tests in the second session, they were partially debriefed about the study. They were told that they might have been shown a slightly different version of the video than the other members in their discussion group. After this partial debriefing,

each participant was asked whether they thought they were in a discussion group in which some other members had seen a different version of the video, and if so to give their reasons.

<center>*Results and Discussion*</center>

Manipulation Awareness Checks

Discussion questions manipulation check. As in Experiment 1, participants in the different-video groups reported their discussions to be just as "influential, informative, open, and free" as the participants in same-video groups (independent t-tests for the four discussion questions, all p's > .05).

Participants were asked what they thought the purpose of the study was. Only one participant guessed correctly and because she was from the different-video group and she was aware of the manipulation, her data were excluded from all analyses.

Post-debriefing check for awareness of manipulation. A chi-square analysis revealed that there was no association between whether participants *thought* they had been in a different–video group and whether they *were* in fact in a different-video group (χ^2 (1, $N = 63$) = 0.01, $p = .95$). When debriefed, the majority of witnesses (65%) thought that others in their group had seen a different version of the video. Of the participants who were in the same-video groups, 65% incorrectly thought that they were in a different-video group. Of the participants who were in the different-video groups, 66% correctly thought that they were in a different-video group. This suggests that there was some naturally occurring disagreement and confusion during the discussion, and that this was not perceptibly greater in the different-video groups than in the same-video groups.

Memory Accuracy

Recognition questionnaire. The mean scores for memory accuracy are shown in Figure 5. Separate 2 x 2 ANOVAs were conducted to investigate the effects of Discussion Type and Warning on participants' accuracy on *misled* and *non-misled items.*

The analysis revealed that there was a significant effect of Discussion Type on *misled items*, F (1, 60) = 5.52, $p < .05$; partial eta squared = .08, such that participants in the different-video groups were less accurate on misled items ($M = 55.33$, $SD = 17.17$) than participants in same-video groups ($M = 67.06$, $SD = 22.50$). Analyses of the individual misled items also revealed that participants in the different-video group had a higher error rate than those in the same-video groups across all misled items except one (thief tipped jewellery), however only one item on its own (thief was drinking beer) reached significance, χ^2 (1, $N = 65$) = 5.14, $p < .05$. Thus, as in Experiment 1, we have demonstrated memory conformity; participants who discussed the video with people who had seen an alternate version incorporated details from the alternate version into their own accounts. As in Experiment 1, participants in the different-video group were also more accurate in their recall of the *non-misled items*, however, in this case the effect was small and not statistically significant, F (1, 58) = 1.22, $p = .27$; partial eta squared = .02.

As in Experiment 1, the main effect of Warning was not significant for *misled items,* F (1, 60) = 0.22, $p = .64$, nor was there a significant interaction between Discussion Type and Warning on *misled items*, F (1, 60) = 0.10, $p = .76$. This suggests that warnings about potential misinformation do not help to mitigate the co-witness memory conformity effect. The main effect of Warning was also not significant for *non-misled items,* F (1, 58) = 0.32, $p = .57$, nor was there a significant

interaction between Discussion Type and Warning on *non-misled items*, F (1, 58) = 0.46, p = .50.

Free recall. Free recall responses were scored in the same way as Experiment 1. Eleven narratives were scored by three independent judges who were unaware of the conditions to which the participants were assigned. There was high inter-rater reliability on all coding classifications (*r*'s ranged from .70 to 1.00, all *p*'s <.05).

Separate 2 x 2 ANOVAs were conducted to investigate the effects of Discussion Type and Warning on *accurate information* and *misleading postevent information reported*. The mean scores are shown in Figure 6.

The univariate ANOVA investigating the effects of Discussion Type and Warning on *misleading postevent information* revealed a significant effect of Discussion Type, F (1, 59) = 6.55, $p < .05$; partial eta squared = .10. Participants in the different-video group were more likely to report the misleading information (M = 0.40, SD = 0.67) than those in the same video group (M = 0.06, SD = 0.35). There was no effect of Warning on misleading postevent information reported, F (1, 59) = 0.55, p = .46, and no interaction between Discussion Type and Warning, F (1, 59) = 0.01, p = .91.

The univariate ANOVA investigating the effects of Discussion Type and Warning on *accurate information* revealed that there was no main effect of Discussion Type, F (1, 59) = 1.09, p = .30, no main effect of Warning, F (1, 59) = 0.30, p = .59, and no interaction between Discussion Type and Warning, F (1, 59) = 0.34, p = .56.

In sum, analysis of free recall accuracy showed the same pattern of results as the recognition questionnaire for both the Discussion Type and Warning manipulations.

Percentage of Participants Who Reported Misinformation during Free Recall

The accuracy scores were recoded to indicate which participants had, during free recall, recalled at least one item of misinformation. It was found that 29% of participants in the different-video group reported at least one misled item, while only 6% in the same-video group inaccurately reported at least one misled item. A chi-square analysis revealed that there was a significant association between Discussion Type and whether or not participants reported at least one misled item, ($\chi2$ (1, N = 65) = 6.18, p < .05). A similar comparison was also made across the two Warning conditions. It was found that 17% of participants who received a warning and 17% of those who did not receive a warning reported at least one misled item. A chi-square analysis revealed that there was no association between whether or not the participants reported misinformation and whether or not they received a warning ($\chi2$ (1, N = 65) = 0.003, p = .96).

Source Monitoring Judgments

As with our analysis of Remember/ Know judgements in Experiment 1, our analysis of source monitoring judgements looked only at data from participants in the different-video group and compared non-misled and misled participants (i.e., those who saw the items in their version of the video vs. those who did not). The number of times each participant reported they remembered the item from the "video only," "discussion only," or "both the video and the discussion" were calculated for both misled items and non-misled items.

A 2 x 2 MANOVA was conducted in which the effects of Participant Type (misled, non-misled) and Warning on source-monitoring (misled item video only, misled item discussion only, misled item both, non-misled video only, non-misled discussion only, non-misled both) were analysed. The analysis revealed that there was

no main effect of Participant Type (F (6, 22) = 3.38, p = .06), or Warning (F (6, 22) = 1.55, p = .21), and no significant interaction between Participant Type and Warning (F (6, 22) = 1.98, p = .11) on the combined dependent variable of source monitoring. This suggests, contrary to our prediction, that warning participants did not improve their source monitoring judgments compared to when no warning was given.

Given that the effect of Participant Type in the above analysis was only marginally non-significant, a second analysis was conducted that omitted the factor, Warning. A one-way MANOVA revealed that there was a significant main effect of Participant Type on the combined dependent variable of source monitoring (F (6, 24) = 2.98, p < .05; Wilks' Lambda = .6; partial eta squared = .43). There was a significant main effect of Participant Type on *misled both* (using a Bonferroni adjusted alpha level of .008, F (1, 29) = 9.75, p < .008) such that non-misled participants were more likely to report that they remembered the items from both the discussion and the video (M = 1.80, SD = 1.15) than were misled participants (M = 0.75, SD = 0.68). There were no differences for the other dependent variables (see Table 2).

This result provides some support for the hypothesis that source monitoring judgments can be used to distinguish accurate information from misinformation. However, its utility is limited given that there were no significant differences regarding how often the misled and non-misled participants attributed information to either the video alone or the discussion alone. This suggests that misled participants did not systematically attribute misinformation to either source, and these participants may simply have had difficulty monitoring the source of their information, despite explicit instructions to do so. In this experiment the final recognition and source monitoring test was preceded by a free recall test for which participants were not

required to monitor their sources. This prior recall may have induced the source confusion (see Roediger, Jacoby, & McDermott, 1996). However, Meade & Roediger (Experiment 2, 2002) also found that participants had trouble identifying information sources following co-witness discussion when the recognition test was not confounded by prior recall. The implications of these findings are examined in more detail in the General Discussion section.

General Discussion

The finding that participants report errant co-witness information when later tested individually supports previous studies investigating the memory of individuals who saw different versions of an event to their co-witness (e.g., Gabbert, Memon, & Allan, 2003; Garry et al., 2008; Paterson, Kemp, Ng, 2009; Wright, Self, & Justice, 2000). The memory conformity effect was robust in both experiments, manifesting in responses to both recognition and free recall questionnaires. Furthermore, the current studies demonstrated that memory conformity can occur in groups in which there was another member who had witnessed the same version of the video as the participant. Research shows that conformity decreases dramatically (by almost 80%) if just one person agrees with the participant (e.g., Asch, 1952). During their discussion about the video, our participants were always in the company of another participant who had witnessed the same version of the video and who was likely to have agreed with their recollection of the event. The fact that memory conformity was still evident in both experiments despite this arrangement only serves to emphasise the power of co-witness discussion to change memory for an event. The memory conformity effect was also resistant to warnings designed to reduce the reporting of misinformation. Remember/know judgements and source monitoring were found to be of only limited value when trying to differentiate memories of events from memories influenced by

co-witnesses. These observations are especially important given that particular attention was given to designing this study to ensure a high level of ecological validity.

In both experiments, it was found that participants reported information provided by their co-witness despite warnings not to do so. This suggests that they may be unable to distinguish what they actually witnessed from the information they discussed with the co-witness. This supports the notion that memory change may be responsible for the conformity observed. Other research on co-witness discussion has found that warnings could significantly reduce, but not eliminate the deleterious effects of co-witness discussion (Meade & Roediger, 2002). The discrepancy between the results of the current study and those of Meade and Roediger may be due to the fact that participants in the current study were given a longer delay between the presentation of postevent information and the warning. Source misattribution is more likely to occur after a longer delay (Underwood & Pezdek, 1998) and therefore the warnings may not be as effective after a longer delay because participants are less able to accurately attribute the source of a memory. With the resultant decrease in effect size, we may simply have not had enough power to detect the effect of warnings. Thus, we cannot completely rule out the role of informational influence in co-witness memory conformity. However, it should be noted that there was no negative effect of warning. This suggests that when working with real eyewitnesses there is probably no harm in using warnings, especially since the cost of doing so is trivial. The practical implication of this result is that warning witnesses about potential contamination of their accounts may have little real utility, especially if the warning is delivered after more than a short period after the event, as is often the case in real eyewitness scenarios. The legal system should therefore not rely too heavily on

such warnings. For example, when a court warns a witness not to provide hearsay evidence, the witness may still do so unintentionally.

A second technique that we employed to establish whether it is possible to distinguish true memories from co-witness misinformation was the use of source monitoring instructions. Our finding that memory conformity occurs even when participants are asked to monitor the source of the memory is consistent with findings from Meade and Roediger (2002) and Gabbert et al. (2007) and is particularly notable because prior research has shown that source monitoring can eliminate the misinformation effect when postevent information is encountered in other ways (Lindsay & Johnson, 1989). Thus, it appears that when postevent information is encountered through postevent narratives and leading questions, the misinformation effect is eliminated if participants are asked to introspect about the source of the information, but when the misinformation is encountered through co-witness discussion, this instruction has no effect on the amount of misinformation reported. Furthermore, Meade & Roediger (Experiment 4, 2002) found that participants were more likely to misattribute the source of misinformation to the original stimulus when it was encountered through co-witness discussion than when it was encountered through the implied social presence of another.

Even in combination with warnings, source monitoring did not eliminate the effect of co-witness memory conformity. This confirms previous research suggesting that co-witness misinformation has an especially powerful influence on memory (Gabbert, Memon & Wright, 2004; Paterson & Kemp, 2006b) and is resilient to attempts to eradicate it. As Meade and Roediger concluded, "it is noteworthy that social contagion is a powerful enough phenomenon to persist even under these stringent testing conditions," (Meade & Roediger, 2002, p. 1007).

There are several reasons why exposure to misinformation through co-witness discussion may be more disruptive to participants' attempts to monitor the source of their memories than other channels for misinformation. First, co-witness discussion may be more easily confused with the actual event because the two sources are more similar than, for example, a written narrative would be to the event. That is, the original event and the discussion of the event share several characteristics in common, including the presence of the co-witness. Furthermore, both the event and the discussion involve the presentation of information visually and verbally. This is in contrast to the presentation of misinformation in a written transcript which shares few characteristics with the original event. A further explanation may lie in the fact that participants may view the co-witnesses as a more credible information source than the experimenter because the participant believes the co-witness saw the same thing as s/he did. People generally assume that information exchanged during the course of a discussion is truthful and accurate (Grice, 1975). The participant has no reason to think that this is not also true of their discussion with the co-witness, and as a result may not make an effort to monitor their sources carefully (Echterhoff, Hirst, & Groll, 2001). When people do not make an effort to monitor the source of their information at the time of encoding because they believe the source to be credible, then presumably they would find it very difficult to separate the sources later on.

A further interview technique that was investigated was the use of remember/know judgments (Tulving, 1985). In Experiment 1 it was found that it may be possible to reduce the amount of co-witness misinformation reported by instructing witnesses to only report information they "remember" and not report information that they just feel they "know". It may be useful to employ this technique to help differentiate between real and co-witness induced memories. Whilst it may have low

face validity, Experiment 1 has demonstrated that this technique has modest predictive power for memory accuracy.

An unexpected finding from Experiment 1 was that participants in the different-video groups gave more accurate accounts for non-misled items than those in the same-video groups and the no-discussion controls. Although this advantage for participants in the different-video groups on non-misled items was not replicated in Experiment 2, the means followed the same pattern.

We can think of several possible (and mutually compatible) explanations for this. Firstly, disagreement in groups can generate better discussion in which more information is elicited from the other group members. Research on "devil's advocates" has shown that group discussion can be enhanced when a member voices a differing opinion (e.g., Janis, 1972). Secondly, when faced with differing information, participants may have steered the group discussion away from topics on which people did not agree (i.e., misled items), and so spent more time discussing the non-misled items compared to participants in the same-video and control groups. Research on the 'collective information sharing bias' suggests that there is a tendency for group discussions to focus on the reiteration of information that group members share in common, with little attention being paid to the exposure of information that is unique to individual members (Wittenbaum, Hubbell, & Zuckerman, 1999). Finally, when faced with conflicting information, group members may have made a greater effort to recall the details of the event. This greater recall effort may have included the use of strategies such as context reinstatement which can improve recall. Future research should explore these possible explanations and further investigate the use of natural discussion groups, rather than no discussion controls.

In sum, the robust memory conformity effect demonstrated in both experiments may be attributable to memory change. This claim is supported in that: a) participants were tested individually and therefore it is unlikely that normative social influence caused the conformity; b) participants reported co-witness misinformation in free recall as well as closed questions so it cannot be a result of demand characteristics of the recall questionnaire; c) memory conformity occurred when participants were encouraged to report what they remembered from the original stimulus and were warned about possible misinformation; and d) memory conformity occurred on a source monitoring test that explicitly called participants' attention to the various possible sources of information. These results are particularly important given that to preserve ecological validity, we used video stimuli rather than photographs, natural discussion with a genuine co-witness, and a relatively long delay. This means that it is likely the results will generalize to real eyewitness scenarios.

References

Asch, S. E. (1952). *Social psychology*: NY: Prentice-Hall.

Azad, T., Lindsay, D. S. & Brimacombe, C. A. E. (March, 2010). Memory conformity: Do people falsely "remember," "know," or "guess" unseen details they learned from another person during co-witness discussion? Paper presented at the American Psychology-Law Society Annual Conference, Vancouver, Canada.

Basden, B. H., Basden, D. R., & Henry, S. (2000). Costs and benefits of collaborative remembering. *Applied Cognitive Psychology, 14,* 497-507.

Belli, R. F. (1989). Influences of misleading postevent information: Misinformation interference and acceptance. *Journal of Experimental Psychology: General, 118,* 72-85.

Betz, A. L., Skowronski, J. J., & Ostrom, T. M. (1996). Shared realities: Social influence and stimulus memory. *Social Cognition, 14,* 113-140.

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influence upon individual judgment. *Journal of Abnormal and Social Psychology, 51,* 629-636.

Forbes, J. R. (2003). The Laws of Australia, Testimony 16.4 [66].

Gabbert, F., Memon, A., & Allan, K. (2003). Memory conformity: Can eyewitnesses influence each other's memories for an event? *Applied Cognitive Psychology, 17,* 533-543.

Gabbert, F., Memon, A., Allan, K. & Wright, D.B. (2004). Say it to my face: Examining the effects of socially encountered misinformation. *Legal and Criminological Psychology, 9,* 215-227.

Gabbert, F., Memon, A., Wright, D.B. (2007). I saw it for longer than you: The

relationship between perceived encoding duration and memory conformity.
Acta Psychologica, 124, 319-331.

Gallo, D. A, Roberts, M. J., & Seamon, J. G. (1997). Remembering words not
presented in lists: Can we avoid creating false memories? *Psychonomic
Bulletin & Review, 4,* 271-276.

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

Garry, M., French, L., Kinzett, T, & Mori, K. (2008). Eyewitness memory following a
discussion: Using the MORI technique with a Western sample. *Applied
Cognitive Psychology. 22,* 431-439.

Heaton-Armstrong, A. (1987). Police officers' notebooks: Recent developments. *The
Criminal Law Review*, 470-472.

Higham, P.A., & Vokey, John R. (2004). Illusory recollection and dual-process
models of recognition memory. *Quarterly Journal of Experimental
Psychology: Human Experimental Psychology,   57,* 714-744.

Hoffman, H. G., Granhag, P. A., See, S. T. K., & Loftus, E. F. (2001). Social
influences on reality-monitoring decisions. *Memory & Cognition, 29,* 394-
404.

Hollin, C. R., & Clifford, B. R. (1983). Eyewitness testimony: The effects of
discussion on recall accuracy and agreement. *Journal of Applied Social
Psychology, 13,* 234-244.

Janis, I. L. (1972). *Victims of groupthink.* Boston: Houghton Mifflin.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring.
Psychological Bulletin, 114, 3-28.

Lindsay, D. S., & Johnson, M. K. (1989). The eyewitness suggestibility effect and memory for source. *Memory & Cognition, 17,* 349-358.

McCloskey, M., & Zaragoza, M. (1985). Misleading postevent information and memory for   events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General, 114,* 1-16.

McDermott, K. B., & Roediger, H. L. III. (1998). Attempting to avoid illusory memories: Robust false recognition of associates persists under conditions of explicit warnings and immediate testing. *Journal of Memory & Language, 39,* 508-520.

Meade, M. L., Roediger, H. L., III. (2002). Explorations in the social contagion of memory. *Memory & Cognition, 30,* 995-1009.

Paterson, H. & Kemp, R. (2005). Co-witness discussion: A survey of police officers' attitudes, knowledge, and behaviour. *Psychiatry, Psychology and Law, 12,* 424-434.

Paterson, H. M. & Kemp, R. I. (2006a). Co-witnesses talk: A survey of eyewitness discussion. *Psychology, Crime & Law, 12,* 181-191.

Paterson, H. M. & Kemp, R. I. (2006b). Comparing Methods of Encountering Post-event Information: The Power of Co-witness Suggestion. *Applied Cognitive Psychology, 20,* 1083-1099.

Paterson, H. M. Kemp, R. I. & Ng, J. R. (2009). Combating Co-witness Contamination: Attempting to Decrease the Negative Effects of Discussion on Eyewitness Memory. *Applied Cognitive Psychology.*

Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory & Cognition, 21,* 89-102.

Roediger, H. L., III, Jacoby, J. D., & McDermott, K. B. (1996). Misinformation

effects in recall: Creating false memories through repeated retrieval. *Journal of Memory & Language, 35,* 300-318.

Roediger, H. L. III, & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 21,* 803-814.

Roediger, H. L., III, Meade, M. L., & Bergman, E. T. (2001). Social contagion of memory. *Psychonomic Bulletin & Review, 8*, 365-371.

Schneider, D. M., & Watkins, M. J. (1996). Response conformity in recognition testing. *Psychonomic Bulletin & Review, 3,* 481-485.

Shaw, J. S., III, Garven, S., & Wood, J. M. (1997). Co-witness information can have immediate effects on eyewitness memory reports. *Law & Human Behavior, 21,* 503-523.

Skagerberg, E. M & Wright, D. B. (2008). The prevalence of co-witnesses and co-witness discussions in real eyewitnesses. *Psychology, Crime & Law, 14,* 513-521.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using Multivariate Statistics* (2nd ed.). New York: Harper Collins.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology, 26*, 1-12.

Tversky, B., & Tuchin, M. (1989). A reconciliation of the evidence on eyewitness testimony: Comments on McCloskey and Zaragoza. *Journal of Experimental Psychology: General, 118,* 86-91.

Underwood, G., & Milton, H. (1993). Collusion after a collision: Witnesses' reports of a road accident with and without discussion. *Applied Cognitive Psychology, 7,* 11-22.

Underwood, J., & Pezdek, K. (1998). Memory suggestibility as an example of the

sleeper effect. *Psychonomic Bulletin & Review, 5,* 449-453.

Weldon, M. S., Blair, C., Huebsch, P. D. (2000). Group remembering: Does social loafing underlie collaborative inhibition? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1568-1577.

Wittenbaum, G. M., Hubbel, A., & Zuckerman, C. (1999). Mutual enhancement: Toward an understanding of the collective preference for shared information. Journal of *Personality and Social Psychology, 77,* 967-978.

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

Wright, D. B. & McDaid, A. T. (1996). Comparing system and estimator variables using data from real line-ups. *Applied Cognitive Psychology, 10,* 75-84.

Wright, D. B., Self, G., & Justice, C. (2000). Memory conformity: Exploring misinformation effects when presented by another person. *British Journal of Psychology, 91,* 189-202.

Wright, D. B., Gabbert, F., Memon, A., & London, K. (2008). Changing the criterion for memory conformity in free recall and recognition. *Memory, 16,* 137-148.

Yarmey, A. D. (1992). The effects of dyadic discussion on earwitness recall. Basic & *Applied Social Psychology, 13,* 251-263.

Yarmey, A. D., & Morris, S. (1998). The effects of discussion on eyewitness memory. *Journal of Applied Social Psychology, 28,* 1637-1648.

Zaragoza, M. S., & Koshmider, J. W. (1989). Misled subjects may know more than their performance implies. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 15,* 246-255.

Table 1.

Summary of Differences between Video A and Video B

Version A of the Video	Version B of the Video
Thief tipped jewellery out of jewellery box.	Thief looked through jewellery box.
Thief was drinking beer.	Thief was not seen drinking beer.
Thief stole camera.	Thief didn't steal camera.
Thief sent to fix water problems.	Thief didn't say why he was sent.
Woman is wearing a headband.	Woman was not wearing a headband.
Actor 1 as thief	Actor 2 as thief[1]

[1] The Recognition Questionnaire did not contain any items regarding the discrepancy between the two actors who played the thief because both actors had very similar features (blond hair, blue eyes, medium build etc) and we couldn't write descriptions that would be true for one actor and false for the other.
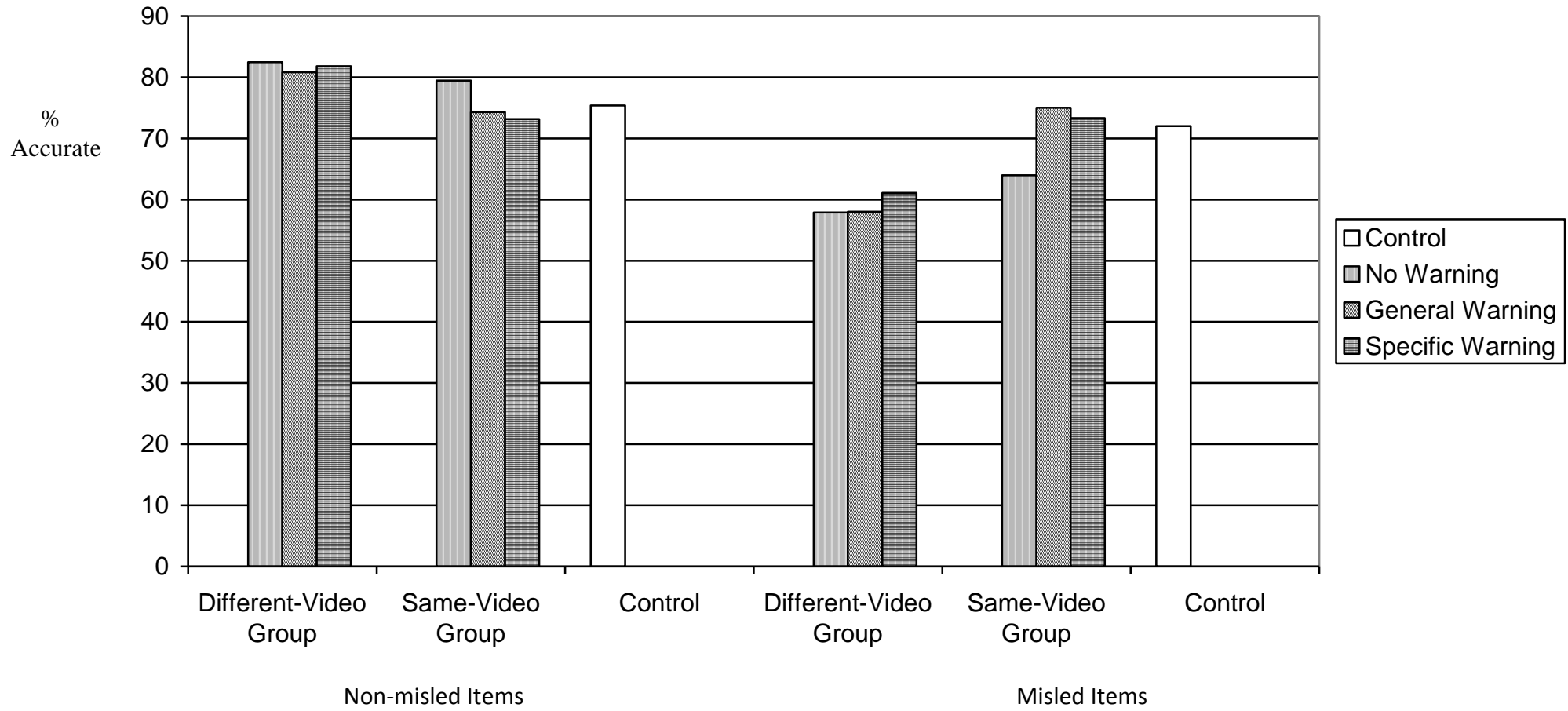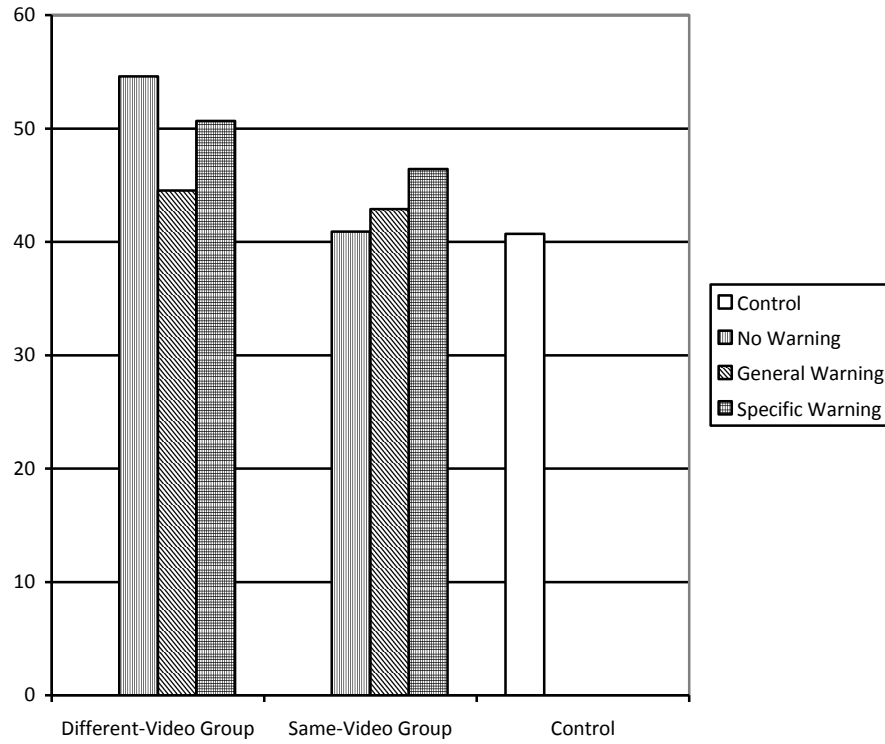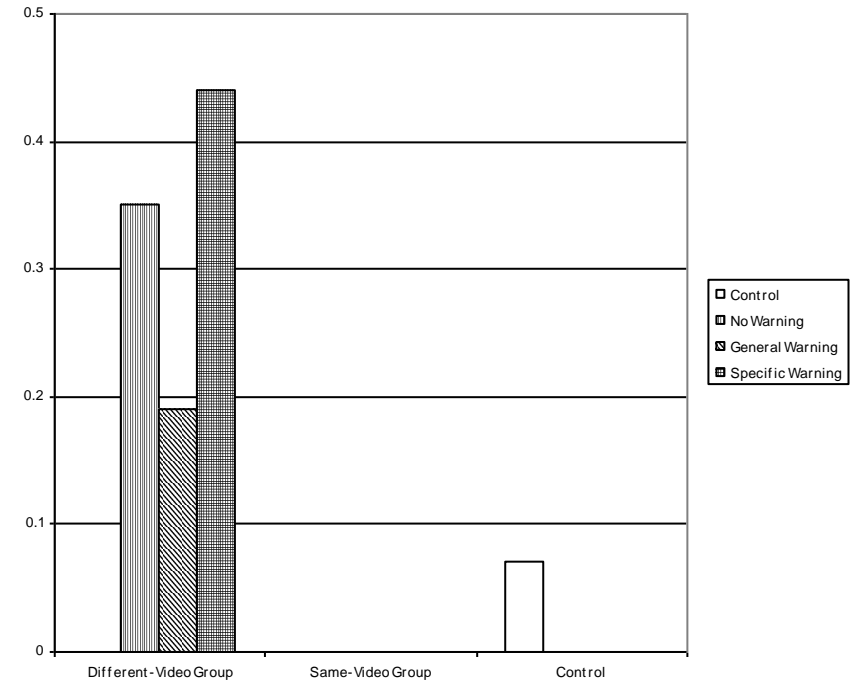
Figure 1. Experiment 1: Recognition Questionnaire accuracy.

Accurate Information Recalled

Misleading Postevent Information Recalled

Figure 2. Experiment 1: Free recall memory accuracy.

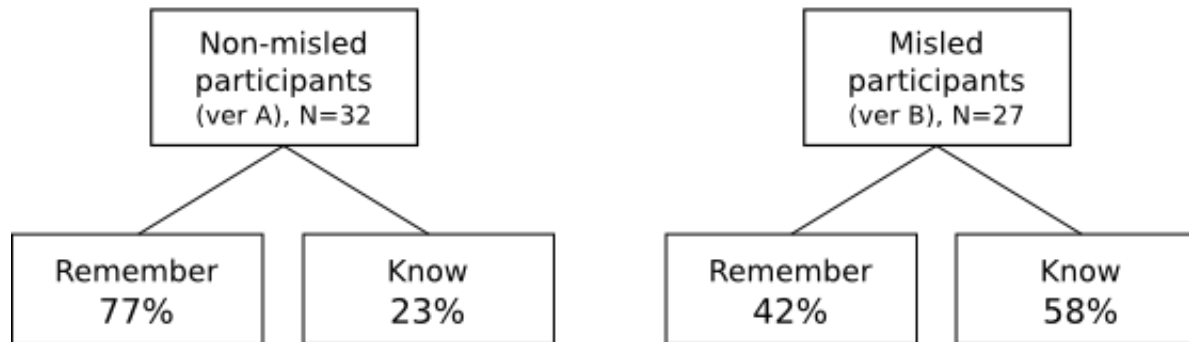## Different Video Groups on items that differed:

```
┌─────────────────┐                    ┌─────────────────┐
│  Non-misled     │                    │    Misled       │
│  participants   │                    │  participants   │
│  (ver A), N=32  │                    │  (ver B), N=27  │
└─────────────────┘                    └─────────────────┘
```

Remember	Know	Remember	Know
77%	23%	42%	58%

Figure 3. Experiment 1: The first analysis of Remember/Know judgments.

Items labelled "true":

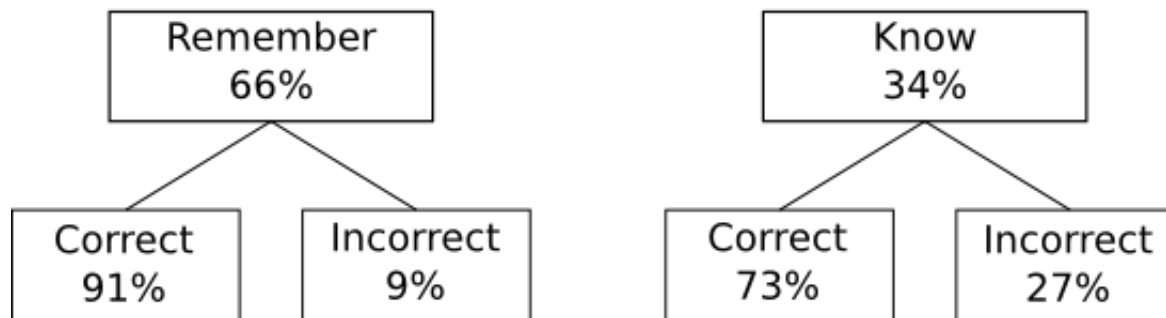| Remember 66%		Know 34%	
Correct 91%	Incorrect 9%	Correct 73%	Incorrect 27%

Figure 4. Experiment 2: The second analysis of Remember/Know judgments. For items that participants said were true, they made a Remember judgment for 66% of those items on average. When making Remember judgments, participants were correct 91% of the time on average. In contrast, when making Know judgments, participants were correct only 73% on average.
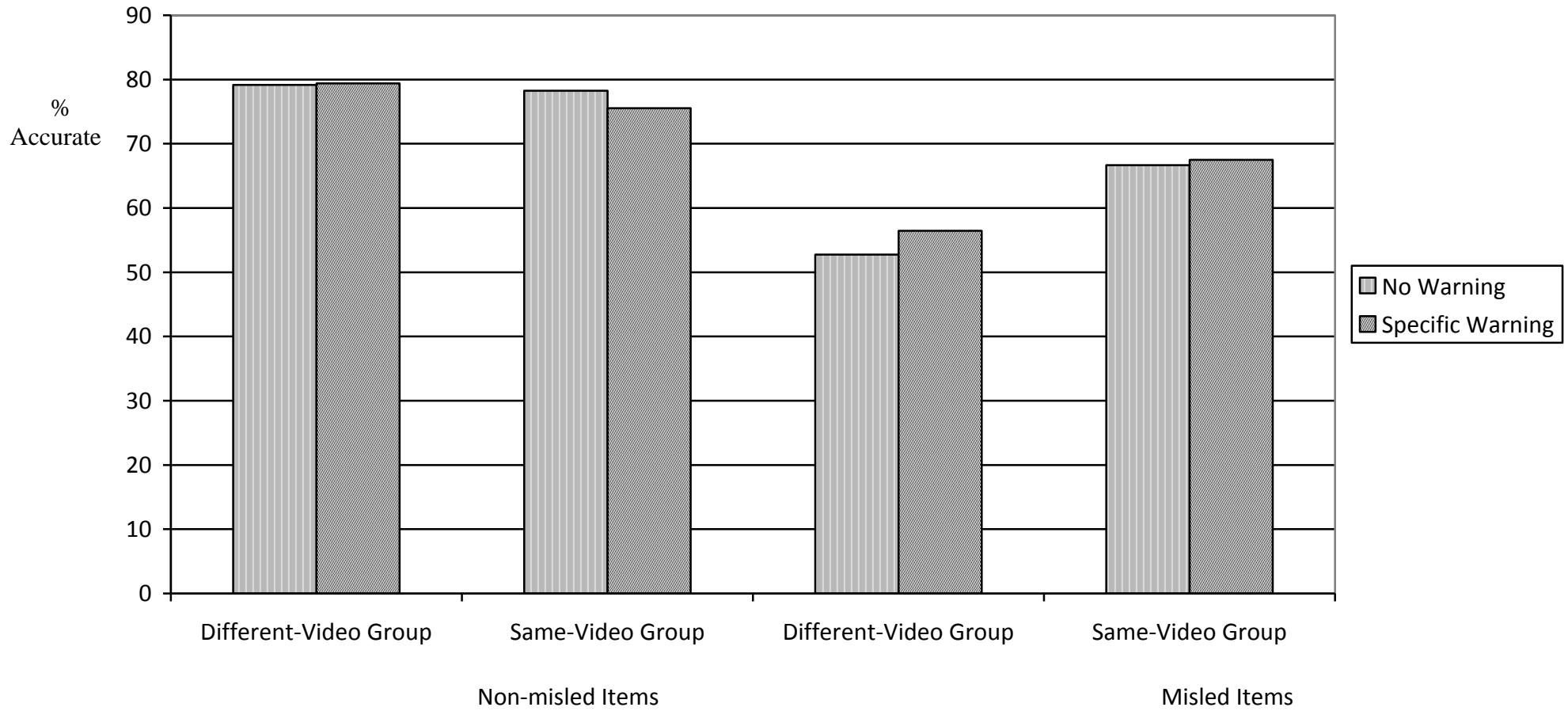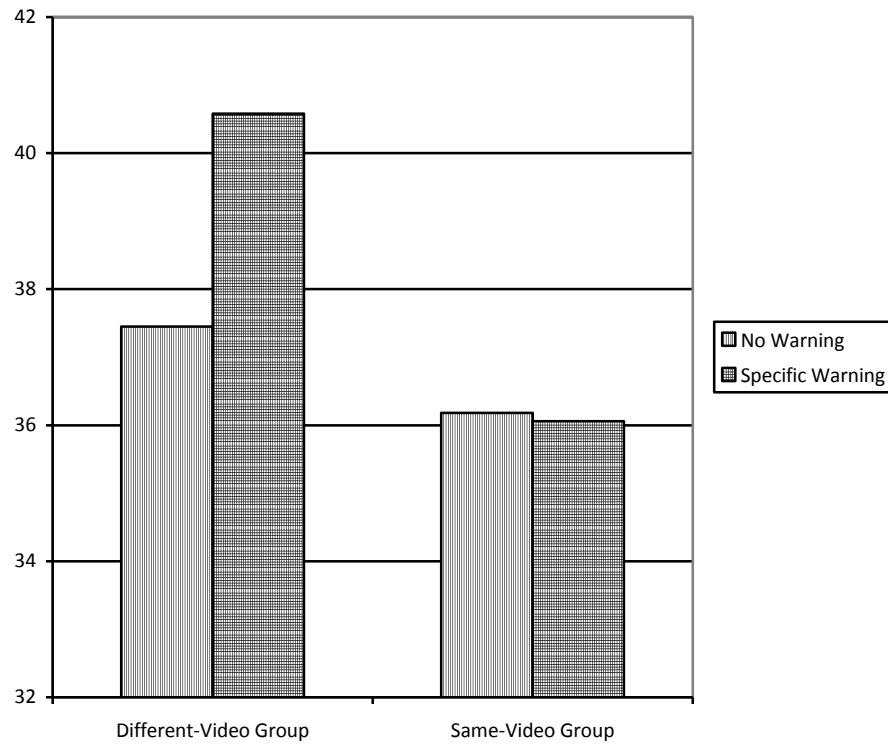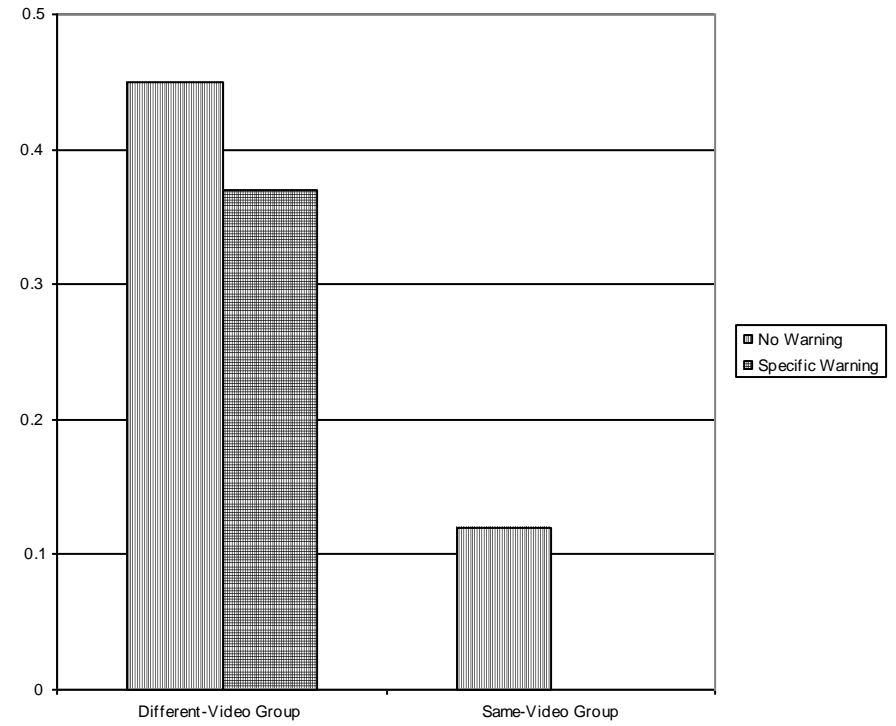
Figure 5. Experiment 2: Recognition Questionnaire accuracy.

Accurate Information Recalled

Misleading Postevent Information Recalled

Figure 6. Experiment 2: Free recall memory accuracy.

Table 2. Experiment 2: Scores for Source Monitoring Judgments

Participant Response		Non-misled Participants		Misled Participants	
		M	*SD*	*M*	*SD*
Misled Items	Video Only	1.07	1.10	1.44	0.81
	Discussion Only	0.33	0.62	0.56	0.63
	Both Video and Discussion	1.80	1.15	0.75	0.68
Non-misled Items	Video Only	5.93	2.58	7.56	3.03
	Discussion Only	0.47	0.74	1.06	1.00
	Both Video and Discussion	5.53	1.81	4.31	2.39