This is an Accepted Manuscript of an article published by Taylor & Francis Group in *Memory*, available online: http://www.tandfonline.com/ 10.1080/09658211.2015.1117641

Why misinformation is reported: Evidence from a warning and a source-monitoring task

Helen Wyler

Margit E. Oswald

Helen Wyler[1], Department of Psychology, University of Bern, Bern, Switzerland;

Margit E. Oswald, Department of Psychology, University of Bern, Bern, Switzerland

[1]Helen Wyler is now at Birmingham City University, Birmingham, United Kingdom.

Correspondence concerning this article should be addressed to
Helen Wyler
Department of Psychology
Birmingham City University
4 Cardigan Street
Birmingham
B4 7BD
United Kingdom
E-Mail: helen.wyler@bcu.ac.uk
Phone: +44 (0)121 331 8531

Author Note

We thank the reviewers for their most helpful and constructive comments and feedback on previous drafts of this manuscript. We also thank Sarah Graf, Yves Linder, and Sandro Salathé for their valuable help in data collection.

Abstract

People report suggested misinformation about a previously witnessed event for manifold reasons, such as social pressure, lack of memory of the original aspect, or a firm belief to remember the misinformation from the witnessed event. In our experiments (*N* = 429), which follow Loftus's paradigm, we tried to disentangle the reasons for reporting a central and a peripheral piece of misinformation in a recognition task by examining (a) the impact a warning about possible misinformation has on the error rate, and (b) whether once reported misinformation was actually attributed to the witnessed event in a later source-monitoring (SM) task. Overall, a misinformation effect was found for both items. The warning strongly reduced the misinformation effect, but only for the central item. In contrast, reports of the peripheral misinformation were correctly attributed to the misinformation source or, at least, ascribed to guesswork much more often than the central ones. As a consequence, after the SM task, the initially higher error rate for the peripheral item was even lower than that of the central item. Results convincingly show that the reasons for reporting misinformation, and correspondingly also the potential to avoid them in legal settings, depend on the centrality of the misinformation.

Keywords: misinformation effect, centrality of misinformation, source monitoring, warning, eyewitness

Why misinformation is reported: Evidence from a warning and a source-monitoring task

For law enforcement authorities, it is crucial that witnesses report their observations at a crime scene as precisely and thoroughly as possible. But testimonies can be distorted, as often-cited examples of far-reaching errors committed by witnesses illustrate. For example, witnesses have unintentionally obstructed criminal investigations by reporting, for instance in the Oklahoma bombing of 1995, that two people rather than one rented a truck or, in the case of the murder of the Swedish foreign minister Anna Lindh in2003, that the perpetrator was wearing a camouflage-patterned military jacket instead of a grey hooded sweatshirt (see Gabbert & Hope, 2013). The probability of such misreports increases under certain conditions. In the 1970s, Loftus and colleagues (e.g. Loftus, Miller, & Burns, 1978) discovered that exposing participants to misinformation about a previously witnessed event can impair their memory reports of the original event. Since then, this finding, which is generally referred to as the *misinformation effect*, has been replicated in many studies and under a range of conditions (see Chrobak & Zaragoza, 2013, Loftus, 2005, and Zaragoza, Belli, & Payment, 2007, for reviews).

The misinformation effect is typically investigated with the classical three-staged paradigm first introduced by Loftus (e.g. Loftus, 1975; Loftus et al., 1978). Participants are first shown an original event such as a theft or a car accident and are later exposed to misleading information about that event. Finally, memory reports are assessed, in the early studies typically with a two-alternative forced-choice recognition test (see also Zaragoza et al., 2007). A misinformation effect can be demonstrated if incorrect answers increase significantly in a group that was exposed to misinformation compared to a control group.

However, such false reports[1] should not be taken as a conclusive evidence of memory impairment or the development of false memory, because the reasons for reporting the suggested misinformation in a recognition task may be manifold (see e.g. Chrobak & Zaragoza, 2013; McCloskey & Zaragoza, 1985). For instance, participants may hold a firm belief that the suggested information really was part of the original event or report the misinformation to comply with the experimenter. Hence, the misinformation effect is probably best understood as a collective term that captures various reasons for false reports.

In the past decades, substantial efforts have been undertaken to better understand the causes of these false reports (see e.g. Chrobak & Zaragoza, 2013, for an overview). A central concern is to determine the extents to which participants' memory is actually impaired and to which they report the misinformation for other reasons. One promising approach to this issue is to examine how different manipulations affect the error rate (e.g. a warning about having possibly been exposed to misleading information; for reviews see Blank & Launay, 2014 and Echterhoff, Hirst, & Hussy, 2005). Another is to examine the effect different types of memory tests have on the error rate (e.g. a source-monitoring task; see e.g. Bodner, Musch, & Azad, 2009; Lindsay & Johnson, 1989; Zaragoza & Lane, 1994). In this article, we argue that combining these two approaches holds great promise for differentiating between different causes of false reports, which would not be possible by examining each approach separately. Within this combined procedure (see Figure 1 for an overview), some participants are warned that they may have been exposed to misinformation, whereas others are not. In addition, all participants solve a post-recognition source-monitoring (SM) task in which they are asked to

[1] In this article, *false reports* always refer to participants reporting misinformation that was suggested to them in the post-event information.

indicate the source on which their answer in the recognition task was based. Please note that the combination of a warning and an SM task has previously been investigated by Zaragoza and Lane (1994) and Higham (1998). In these studies, however, the warning was not manipulated between subjects and participants did not answer a recognition task before the SM task. The warning was used to eliminate task demands and response biases, yet both studies still reported a misinformation effect.

Figure 1

In the following, we first elaborate on different causes that may lead people to report misinformation in a recognition task. Next, we elucidate what a warning may contribute to the discrimination of these causes and finally discuss the additional benefits expected from the SM task.

Reasons for False Reports in a Recognition Task

While accomplishing a recognition task, misinformed participants find themselves in a range of different situations, from which many reasons for reporting misinformation may arise. The participants may, for instance, differ in the completeness of their memory of the original information, their awareness that they received contradictory information about the original event, their reaction to perceived social pressure or persuasion, and the depth of their memory retrieval. Based on the pivotal work of McCloskey and Zaragoza (1985) as well as on those of Belli (1989), Blank (1998, 2009), and Chrobak and Zaragoza (2013), we outline four different categories of reasons for false reports. Although these categories may overlap to some extent, thus presenting difficulties in assigning some false reports unequivocally to one particular category, they are suitable for our purpose.

Deliberation

Deliberation (Belli, 1989) refers to instances in which participants decide to report misinformation although they know or at least suspect that the post-event information (PEI) does not correspond with their own memory of the original event.[2] On the one hand, participants may want to comply with the substantial task demands of the recognition task (see e.g. Lindsay, 1990). On the other hand, they are usually motivated to answer correctly. In many studies, the experimenter, who participants assume is knowledgeable about the original event, provides the PEI. Consequently, participants presume the PEI to be consistent with the original event, unless they suspect an attempt at deception or are warned explicitly (*consistency assumption*, see e.g. Blank, 2009) and report the misinformation because they believe it to be more reliable than their own memory. For the same reason, participants are increasingly likely to accept a suggested item if they think the source that provided the PEI (e.g. another participant) has more reliable information than they do (*informational influence*; Deutsch & Gerard, 1955; e.g. Allan, Midjord, Martin, & Gabbert, 2012; Horry, Palmer, Sexton, & Brewer, 2012; for a recent review, see Gabbert & Hope, 2013).

Recency Bias

In many situations, participants do not carefully check exactly what they remember from the original event and/or the source from which they remember a specific piece of information. Depending on participants' understanding of the situation and the task

---

[2] Of course it cannot be ruled out that a deliberation error evolves into a false belief (see below) due to imagination inflation, elaboration, or visualisation processes. For instance, Zaragoza and Lane (1994) found that a reactivation of the originally encoded information during processing the misinformation increases source misattributions. This situation may be similar to what happens to participants whilst answering a recognition question.

instructions, they may simply rely on decision criteria such as familiarity or retrieval fluency (see e.g. Blank, 1998; Horry, Colton, & Williamson, 2014) instead of checking their memory critically. This may be the case, for instance, if they *a priori* assume that the information from the original event and the PEI will be consistent, and hence see no reason to reflect on the source of a memory (cf. consistency assumption).

Errors due to a recency bias are characterised by the fact that the original information would be available and a correct source attribution would be possible in principle. However, as long as participants are not motivated to use this information, the misinformation may be at an advantage over the original, for instance because of the misinformation's recency (see e.g. Belli, Lindsay, Gales, & McCarthy, 1994). Errors due to a recency bias are related to deliberation insofar as both the original and the suggested information are principally available in both cases, and that the person reports the suggested item nonetheless. The crucial difference is that in the case of a recency bias participants do not realise at the time of the memory test that they report a piece of information they only remember from the PEI.

Best-Guess Error

In the cases of both deliberation and the recency bias, it is assumed that participants principally have some memory of the original item. However, there may be situations in which participants entirely lack memory of a specific detail of the original event, either because they never encoded it, or because they forgot it due to normal memory processes (McCloskey & Zaragoza, 1985; see also e.g. Belli & Loftus, 1996). This situation is especially likely to occur if the recognition question concerns a peripheral aspect to which little or no attention was paid whilst observing the original event. If participants are aware of having no memory of the original information but do remember the misinformation, they most probably report this as it is the only information at their disposal; it is their *best guess*[3].

False Belief

False beliefs describe the situation in which participants report the misinformation in a firm belief that it actually occurred, either with or without recollecting specific details of it. This category contains several different mechanisms, which have in common that the resulting errors cannot be prevented or undone by measures taken at the time of the memory test. For example, the misinformation may impair the memory of the original event, either by altering the original memory trace or by making it more or less inaccessible, so that participants mistakenly believe that they remember the suggested item from the original event (see e.g. Loftus, 2005). It should be noted, though, that people might commit a false belief error in one situation but not in another. For instance, the error might not appear if additional retrieval cues provide access to the original information (e.g. Frost & Weaver, 1997) or if the participant is highly motivated. Thus, whether an error is categorised as a false belief to some extent depends on situational factors.

Effects of a Post-Warning

Evidence about the error-reducing effects of a warning given after the PEI but prior to the memory test (*post-warning*) is mixed (see e.g. Echterhoff et al., 2005). Yet the results

[3] The notion of what we describe as *best guess error* has previously been described by McCloskey and Zaragoza (1985) although the authors did not use a specific term for this type of error. Belli later coined the term *misinformation acceptance*. We did not use this term, as we think it is less intuitive than *best guess* and could erroneously be interpreted as implying that participants accept the misinformation against their better knowledge (cf. deliberation).

reported by Blank and Launay (2014) in their recent meta-analysis suggest that, on average, a post-warning reduces the misinformation effect to less than half of its original size, although it has to be noted that effect sizes varied considerably and also partly depended on the type of warning. Investigating the effects of post-warnings (in the following also just called warning) is expected to help in distinguishing between different reasons for false reports. Such warnings are assumed to reduce or even eliminate task demands and informational influence, thereby preventing deliberation errors. Furthermore, they raise participants' motivation to monitor the source of information carefully, as one of the two sources, the PEI, is declared to be potentially unreliable, which counteracts the recency bias. By contrast, false beliefs and best-guess errors are expected to occur in spite of a warning. A false belief persists by definition, and the best-guess error persists because a participant can only benefit from a warning if they have at least some memory of the original item (Christiaansen & Ochalek, 1983). As participants do not remember the original item in the case of a best-guess error, they might continue relying on the PEI unless they receive a very strong warning that informs them that every single piece of information from the PEI was incorrect (logic of opposition; Lindsay, 1990).

Benefits of an Additional Source Monitoring Task

In an SM task, participants are asked to specify the source from which they remember a certain piece of information, and they typically receive the following four options for an answer: *original event only*, *post-event information only*, *both*, and *neither* (see Lindsay & Johnson, 1989). Whereas in a recognition task the misinformation is solely accepted as the correct answer to a question about the original event, in an SM task participants have to misattribute the suggested information explicitly to the original event in order to commit an error. Hence, the SM task can be conceived as a conservative measure of the misinformation effect, whereas the recognition task can be understood as a more liberal measure. Empirical evidence suggests that the error rate is indeed lower in an SM than in a recognition task (see e.g. Lindsay & Johnson, 1989; Zaragoza & Koshmider, 1989; Zaragoza & Lane, 1994).

Somewhat differently from other studies, we used a post-recognition SM task to check whether those participants who gave incorrect answers in the preceding recognition task (i.e. who chose the misinformation) actually attribute their answers to the original event. As participants may simply guess when answering the recognition task (i.e. their recognition answer is pure guesswork), our fourth response option in the SM task was *guessed* (instead of *neither*). Note that guessing may also play a role in the SM task itself (Blank 1998). Participants may be unsure of the actual source from which they remember a piece of information but are obliged to select one of the four response alternatives (see also Higham, 1998). Yet, barring those cases, participants' answers in the SM task indicate the type of error. An affirmation of a false report by attributing it to the original event would be expected in the case of false beliefs, but not if a false report is due to a best guess. In the latter case, the incorrect answer will be attributed to the PEI by definition. Furthermore, false reports due to deliberation or the recency bias should be adjusted, too. For the former, the *post-event information only* option may inform participants that some information could have been contained in the PEI only, while for the latter, the SM task should inherently promote a thorough source monitoring. Table 1 summarises theoretical assumptions about the effects of a warning and participants' answers in the SM task. The crucial point is that a warning allows errors due to deliberation or the recency bias to be identified, but the remaining errors can be differentiated into best-guess errors and false belief only by an additional SM task.

Table 1

This Study

Overall, three experiments and one post-test were conducted to explore the potential of combining a warning (manipulated between subjects) and a post-recognition SM task to better understand the nature of false reports that occur in a recognition task (see also Figure 1). To test this approach, participants were given misinformation on two aspects of the original event that differed in their centrality. This is of special interest because, although the misinformation effect is often found to be more pronounced for peripheral than for central items (Dalton & Daneman, 2006; Echterhoff, Groll, & Hirst, 2007; Heath & Erickson, 1998; Roos af Hjelmsäter, Granhag, Strömwall, & Memon, 2008; Roebers & McConkey, 2003; Shapiro, Blackford, & Chen, 2005; Wright & Stroud, 1998; but see e.g. Luna & Migueles, 2005, and 2009), little is known about the underlying causes of this difference.

In general, people have been found to remember central aspects of the original event better than peripheral ones, especially if the event is emotional (for a review, see Christianson, 1992). As Ibabe and Sporer (2004) argue, this finding is most likely caused by people directing their attention more to central than to peripheral aspects of an event. Thus, they can be assumed to have a weaker if any memory of peripheral aspects. On the one hand, this may result in a larger number of best-guess errors, because participants are aware that they did not pay attention to this aspect of the original event and that they rely on information they remember only from the PEI (McCloskey & Zaragoza, 1985; see also Dalton & Daneman, 2006; Roos af Hjelmsäter et al., 2008). On the other hand, however, one could also argue that the weak memory may foster false beliefs, as it is more difficult for the participant to detect a discrepancy between the original event and the PEI. As a result, people may readily integrate the misinformation into their memory (see e.g. Tousignant, Hall, & Loftus, 1986). For a central item, by contrast, best-guess errors are less probable, as people are more likely to encode and remember the critical aspect of the original event. Yet, for instance errors due to deliberation cannot be excluded with certainty. To shed light on the role of the different errors, we investigated two main research questions.

First, we were interested in whether the occurrence of recognition errors differs between the three conditions realised in our experiments (misinformed and unwarned, misinformed and warned, control). More specifically, with the recognition task we examined (a) whether the misinformation effect found in many prior studies (see e.g. Loftus, 2005) can be replicated, and (b) whether the warning reduces or even eliminates the occurrence of false reports. Both sub-questions were analysed separately for the two items. The PEI may have a stronger influence in the recognition task if the suggested item is peripheral rather than central, because peripheral aspects of the original event are usually remembered less, and are thus more prone to the different kinds of errors, as for instance the best-guess error. The warning is expected to primarily avert deliberation or the recency bias, but to have little effect on false beliefs or best guesses. Thus, it is of interest to examine whether misinformation and a warning have different effects on a peripheral and a central item.

Second, we tested whether the effects that we found in the recognition task persisted in the SM task. The SM task allows participants to revise incorrect answers in the recognition task either by indicating that the reported information was remembered only from the PEI or by acknowledging that the answer was pure guesswork. Specifically, we were interested in whether participants in the misinformed and warned group revised their incorrect recognition answer less than those in the misinformed and unwarned group.[4] This could happen because warned participants were already motivated to elaborate their memory retrieval (see Echterhoff et al., 2005), were reassured that their memory probably was the correct one, and/or felt less obliged to follow the demands of the experiment. Thus, while participants without a warning had the opportunity at this juncture to revise their false report in the SM task, participants with a warning might have avoided a false report from the very beginning. Therefore, after the additional SM task, the error rate of participants without a warning could be almost the same as that in the warned group. The question whether errors in the recognition task persist in the SM task will also be analysed separately for both the peripheral item and the central item. If false reports about a peripheral (compared to a central) item are due more to best-guess errors, false reports about the peripheral item are more likely to be revised in the additional SM task than false reports about the central item. Thus, the error rate

[4] Because the error rate in the SM task depended on that in the recognition task, we could not rule out that the recognition task and the SM task interacted in some way. Therefore, the error rates in the recognition task and in the SM task were not compared directly. Instead, we assessed the question whether misinformed participants with or without a warning benefitted to a similar extent from the opportunity to amend an incorrect answer indirectly by examining whether the relation between the two groups changed from the recognition task to the SM task. If participants benefit from the SM task to the same extent in both groups, the relative error rates will remain similar; if not, a possible effect in the recognition task may disappear or even be reversed in the SM task. The same applies to the question whether adjustments are equally likely in the peripheral and the central item.

for the peripheral item may be almost at the same level as that for the central item after the additional SM task.

Working with only two items allowed us to provide misinformation on a very central item without overly arousing participants' suspicion. As a consequence, however, data analyses were based on frequencies, which requires rather large samples to achieve satisfying power. Therefore, we ran replications with minor variations between the experiments. The overall analyses are reported in the result section. Detailed information on the individual experiments and the individual analyses can be found in the supplemental online material.

Method

Cover Story

The cover story had to meet two requirements. First, for reasons of ecological validity it could not focus on learning or memory. Whilst observing a crime, witnesses often do not give their undivided attention to the critical event. They may either be distracted by their own feelings (e.g. fear) and thoughts (e.g. where to look for help; see also Lane, 2006) or may not even be aware that they have just witnessed something important. If participants have the expectation of a subsequent memory test, however, they might process the original event and the PEI in a depth that they would not be able to reach in a real-life setting. Second, we did not want participants to attribute the PEI to the experimenter, so as to reduce both task demands and the problems associated with the consistency assumption (see e.g. Blank, 1998, 2009).

The experiment was announced as a study on computer-based brainstorming about situation-related crime prevention[5]. Participants were told that their brainstorming partner was a previous participant, and that they would be provided with information (e.g. suggestions for prevention measures) generated by this previous participant. They were further informed that their own suggestions for prevention measures would be placed at the disposal of a subsequent participant. The original event was presented as an example case video, which ostensibly was part of the information that participants received on crime prevention at the beginning of the study. To introduce the PEI, participants were informed that one aim of the study was to investigate the effects that a prior exchange of information between the participants has on the efficiency of brainstorming. The misinformation was embedded in this exchange of information.

Materials

Original event. In the experiments, the original event was an 80-second clip of a robbery filmed from a CCTV perspective. After taking money from a cash dispenser, the eventual victim (a man in his late twenties) was approached by one member of a group of three young men (perpetrator 1). Perpetrator 1 pretended that he wanted to exchange a high-denomination note for some smaller ones. As the victim took out his wallet, perpetrator 1 tried to snatch it but failed. Thereafter, perpetrator 1 and a second man from the group forced the victim towards a wall. The second man took the wallet from the victim and left the scene.

[5] The topic of crime prevention accounted well for why a video of a robbery (original event) was shown. Situation-related crime prevention was chosen to prevent participants from focussing too narrowly on violence in general or on the behaviour of the perpetrators during the brainstorming, which otherwise might have promoted critical intrusions.

Perpetrator 1 punched the victim in the stomach, which made the victim fall to the ground, insulted him, and thereafter fled. The third man from the group (passive perpetrator) did not intervene in the incident at all and was the first to leave the scene. Two women, who had witnessed the whole incident, took care of the victim and called the police as soon as all the perpetrators had left.

Selection of the critical items. To explore our research questions, we selected one peripheral and one central aspect of the original event about which participants were to be misinformed. Although there is no set definition, for study purposes centrality is most often conceptualised either in terms of an item's importance for the event (e.g. Dalton & Daneman, 2006; Heath & Erickson, 1998; Roos af Hjelmsäter et al., 2008) or in terms of its location/salience (e.g. Belli, Windschitl, McCarthy, & Winfrey, 1992; Christianson & Loftus, 1991, cited in Ibabe & Sporer, 2004). We defined centrality in terms of an item's importance for the event. The two critical aspects were selected from a pool of 22 aspects (see Table A1 for an overview), which were later converted to the 22 memory questions used in the experiments. Twenty of these questions, which refer to aspects of the original event about which no misinformation is provided, served as our *comparison items*. The remaining two questions are the *critical items*, which cover the two critical aspects on which participants are misinformed in the PEI.

Two pretests were run to examine whether our two critical aspects (central: how the victim was treated when lying on the floor; peripheral: headwear of the passive perpetrator) met the following two requirements. Firstly, participants should indeed consider the central aspect to be more important to the original event than the peripheral one. Secondly, at least some participants had to accept the misinformation provided about the central aspect to avoid floor effects with the central item. Pretest 1 assessed the perceived centrality of the 22 aspects. Participants (*N* = 13) were told that they were going to watch a video (original event), which would be used in a future study. This future study would examine the generation of ideas about how to prevent certain crimes, and the video at hand was to be shown as an example of such a crime. After the participants in the pretest had watched the video, they were asked (a) to rate the importance of the 22 aspects in the light of this information, and (b) to indicate how much attention they had paid to these aspects (both 7-point Likert scales). The importance and the attention ratings were highly correlated, $r_s = .85$, $p < .001$. Hence, the two ratings were combined into one centrality index. The central critical aspect (*Mdn* = 5.00) was rated as being significantly more central than the peripheral one (*Mdn* = 2.00), Wilcoxon test, $z = 2.80$, $p = .005$ (two-tailed), $r = 0.78$. Hence, the results show that participants indeed perceive the peripheral item to be significantly less central. A median split was performed on the 20 remaining aspects to create two groups of items for later analyses, one of them intended to be the comparison group for the central item (*Mdn* = 4.85; e.g. number of witnesses on site), the other the comparison group for the peripheral item (*Mdn* = 2.05; e.g. shoes worn by perpetrator 1). The allocation of the 22 aspects can be seen from Table A1. Pretest 2 (*N* = 22) confirmed that our second requirement (false reports also occurred for the central item) was met, too. In the experimental group, 45.5% of the participants chose the wrong answer; in the control group, this proportion was 27.3%.

Post-event information. The misinformation on each of the two critical items was embedded in the alleged partner's answers to two different questions concerning the video (overall about 150 words). The answers were based on actual descriptions of the original event gathered in another pretest (*N* = 14). The misinformation on the central aspect was introduced in a question about how the victim was treated by the perpetrators: "*As the victim* *was lying on the ground, the perpetrator kicked him in the stomach*"[6] (in fact, the victim was not touched when lying on the ground). The misinformation on the peripheral aspect was introduced in a question about the appearance of the passive perpetrator: "*This was the guy with the hood pulled over his head*" (in fact, he wore a woollen hat). The misinformation was omitted in the control condition.

Warning. Prior to the memory test, participants in the warned group were informed that it was within the realms of possibility that they had been exposed to misinformation: *"When answering the following questions, please take into account that some of the former participants were instructed to incorporate incorrect information into their answers in the information exchange."* The other groups did not receive this additional information.

Memory test. The first part of the memory test consisted of 22 two-alternative forced-choice recognition questions. Twenty of them related to the comparison items and two to the critical items. The recognition question and the two response alternatives for the central critical item were: *What happened to the victim when he was lying on the ground? A) One of the active perpetrators kicked the victim in the stomach, insulted him, and ran off. B) One of the active perpetrators insulted the victim and ran off.* The question and response alternatives for the peripheral item were: *What did the 'passive' perpetrator wear on his head? A) A woollen hat. B) A hood.* Both the 22 items and the response alternatives were presented in random order. For data analysis, the answers in the recognition test were coded as *correct/incorrect*.

[6] The experiments were conducted in German. Hence, all quotes correspond to our English translations of the original wording.

The second part of the memory test consisted of the SM task. For each item, participants first reread the recognition question and the answer they had selected (e.g. *You were asked "What did the 'passive' perpetrator wear on his head?" and you answered "A hood".*). Subsequently, they were asked to indicate the source(s) of information on which their answer was based. They could choose from four alternatives: *Video, Answers from the partner, Video and answers from the partner,* and *I guessed*. In the SM task, an answer was categorised as an error only if participants chose the suggested answer in the recognition task *and* attributed this incorrect answer to the video (i.e. the SM answer was either *video* or *video and answers from the partner*). All other combinations of recognition and SM answers were understood as not potentially pointing towards the development of a false belief.

Procedure

Participants were tested in groups of five at most and were separated by partition walls. At the beginning of the study, a computer program randomly assigned participants to one of the conditions. All instructions were provided on the computer screen. After having received detailed information on the procedure of the brainstorming study, participants first watched the video (*original event*). Then, the first distractor phase (average duration of 22 minutes) began. Participants read some information on situation-related crime prevention, worked through an additional temporising task, and answered four questions about the video they had seen and the information they had read. The questions did not allude to critical aspects of the video, about which misinformation was provided later. Next, participants read their partner's answers to similar but not identical questions (*PEI*). Subsequently, they worked through another temporising task and completed the brainstorming, the second distractor phase (average duration of 14 minutes). Then, all participants read the instruction for the recognition task, which asked them to answer questions about the original event. They were instructed: "*For simplicity, imagine you were a witness of the incident you saw in the video (case example) and the police were going to question you about the incident.*" Participants in the warned group received the warning that they may have been exposed to misinformation (see p. 18) in addition to this instruction. The recognition task was followed by the SM task. Finally, manipulation check questions and demographic questions were answered. The manipulation checks concerned questions such as how participants assessed their own knowledge of the incident compared to their partner's, whether they remembered the warning, and whether they believed that they had been provided with information actually generated by a former participant (ranging from *1 not at all* to *5 absolutely*). Participants were fully debriefed and thanked for their participation. Undergraduate psychology students participated in exchange for partial course credit; all other students received a soft drink voucher and a chocolate bar. Informed consent was obtained from all participants. All experiments were approved by the ethical review board of our faculty (application no. 2013-3-345853.

Design

The experiments used a mixed design with a three-level between factor (*misinformed and unwarned, misinformed and warned, control*), and a two-level within factor (*central* and *peripheral*). In the control condition participants were neither warned nor exposed to misleading information. Since we worked with frequencies (error rates in both the recognition task and the SM task) rather than with normally distributed interval scale data, we tested the relevant research questions by means of chi square tests.

Rationale for an Overall Analysis

Experiments 2 and 3 were basically replications of Experiment 1. However, some minor changes from experiment to experiment enabled some additional questions to be examined. Detailed information on the underlying rationales for these variations and on the analyses of the individual experiments are presented in the supplemental online material. Overall, the main findings of the experiments were similar, with some exceptions (see also Table B1). First, no misinformation effect was found in Experiment 1. This was due to an unusually high error rate in the control group, which was replicated neither in a post-test (N = 33), which was run to collect further data for Experiment 1, nor in the following experiments. Second, the warning effect observed in Experiments 1 and 2 was not replicated in Experiment 3.

Data analyses that are based on frequencies require rather large samples to achieve satisfying power. Hence, the power of the individual experiments was quite low. For instance, a difference as large as 24% failed to reach significance in Experiment 1. The low power might also explain the inconsistencies in the findings. If these differences are due to sampling errors, it will be more interesting to focus on the overall analysis of the results so as to obtain a more precise estimate of the actual effects. Overall, additional analyses did not reveal systematic differences between the experiments (see the supplemental online material for further information). Therefore, we decided to report the results on the aggregated data from the three experiments and the post-test.

Overall Sample

Overall, 520 participants were tested. Of these, 91 participants had to be excluded due to not remembering the warning (31), problems with the video presentation (10), disturbances during the video presentation or major interruptions during the experiment (15)[7], not

[7] For instance, four participants had to be excluded due to a fire practice, another four because of an interruption while they were watching the video.

believing at all that the information was actually generated by a former participant (25), or knowing the video already (10). The final groups consisted of 174 (misinformed and unwarned), 138 (misinformed and warned), and 117 (control group) participants. In Experiment 3, half of the misinformed participants worked through another type of SM task. Thus, the groups analysed for the additional SM task consisted of 148 (misinformed and unwarned), 114 (misinformed and warned), and 117 (control) participants. Mean age was $M = 22.46$ years ($SD = 3.54$), and 77.2% of the participants were women. Two hundred and twenty-nine (53.4%) participants were undergraduate psychology students; the others were students from other disciplines.

Results

Data from 429 participants were included in the statistical analyses of the recognition task and from 379 in those of the additional SM task. Comparisons of proportions, Cohen's *h* (with small, medium, and large effects at 0.20, 0.50, and 0.80, respectively; Cohen, 1988), and Cramér's phi (with small, medium and large effects at .10, .30, and .50, respectively; Cohen, 1988) were calculated by hand. The statistical software R was used to calculate 95% confidence intervals for the differences in proportions (R package *ExactCIdiff*). Please note that there may be rounding differences, especially for the differences calculated with the R package. All other statistical analyses were conducted with SPSS 21.0. Statistical tests are reported two-tailed, and, because of the multiple testing per research question, a Bonferroni correction was applied, resulting in an alpha level of .017. The critical alpha level for additional analyses was .006.

An overview of the error rates for both the particular experiments and for the aggregated data can be found in Table B1. Table 2 gives overviews of the differences in the error rates between the groups, and summarises the overall findings for the two items and the two measures regarding the questions whether a misinformation effect was found, whether the warning had an error-reducing effect, and whether the warning even eliminated the misinformation effect.

Table 2

Baseline Error

The baseline error refers to the error rate found in the control group, i.e., when participants were not exposed to misinformation. In the recognition task, the baseline error did not significantly differ between the central item (23.9%) and the peripheral item (34.2%), difference = 10.3%, 95% CI [-0.01, 21.1], $z = 1.74$, $p = .082$, $h = 0.30$. In contrast, the error rate tended to be higher for the central item (23.1%) than the peripheral item (11.1%) in the SM task, difference = 12.0%, 95% CI [2.1, 21.1], $z = -2.46$, $p = .014$, $h = -0.49$.

Recognition Task

The chi-square test of independence conducted on the central item revealed a significant association between the three groups and the accuracy of the answer (correct/incorrect), $\chi^2(2, N = 429) = 32.63$, $p < .001$, $\phi_c = .28$. Thus, column proportions of the different groups were compared (critical alpha level = .017). The comparisons revealed a significant misinformation effect, with an increase in the error rate from 23.9% (control group) to 56.9% (misinformed and unwarned group), $z = 6.05$, $p < .001$, $h = 0.69$, and a significant warning effect (decrease from 56.9% to 37.7%, $z = 3.45$, $p = .001$, $h = 0.37$). The difference between the misinformed and warned (37.7%) and the control group (23.9%) was significant, indicating that the misinformation effect was not fully eliminated by the warning, $z = 2.41$, $p = .016$, $h = 0.30$.

The test of independence for the peripheral item was also significant, $\chi^2(2, N = 429) = 52.96$, $p < .001$, $\phi_c = .35$. A significant misinformation effect was observed (increase from

34.2% to 74.7%; $z = 7.39$, $p < .001$, $h = 0.84$), but no warning effect (decrease from 74.7% to 68.8%; $z = 1.14$, $p = .254$, $h = 0.13$). The error rate was significantly higher in the misinformed and warned group (68.8%) than in the control group (34.2%), $z = 5.88$, $p < .001$, $h = 0.71$.

We were additionally interested in whether the effects of the misinformation and the warning differed significantly between the central and the peripheral item. To this end, we had to compare the differences of the differences. Concerning the misinformation effect, we tested whether the differences in the error rates between the control group and the misinformed and unwarned group for the central item (difference = 33.0%) and for the peripheral item (difference = 40.5%) were significantly different from each other. Concerning the effect of the warning, we tested whether the differences in the error rates between the warned and the unwarned groups for the central item (difference = 19.2%) and for the peripheral item (difference = 5.9%) were significantly different from each other. The formula was adjusted to take the partial dependency of the data into account. The misinformation effect was not significantly stronger in the peripheral item than in the central item, difference of differences = 7.6%, $z = 0.98$, $p = .328$. The comparison with respect to the warning also failed to reach significance, difference of differences = 13.3%, $z = 1.76$, $p = .078$.

Considering the two misinformed groups together, participants committed fewer errors with the central item (48.4%) than with the peripheral item (72.1%), difference = 23.7%, 95% CI [15.9, 31.3], $z = 6.24$, $p < .001$, $h = 0.49$. Post-hoc statistical analyses revealed that this held true for the misinformed and warned group as well as the misinformed and unwarned group, difference = 31.2%, 95% CI [19.0, 42.6], $z = 5.46$, $p < .001$, $h = 0.58$ and difference = 17.8%, 95% CI [7.8, 27.7], $z = 3.57$, $p < .001$, $h = 0.41$, respectively.

Source-Monitoring Task

An error in the SM task occurred if participants misattributed the misinformation they had reported in the recognition task either to the *video* or to the *video and answers from the partner*. A significant association between the groups and the central item was also observed in the SM task, $\chi^2(2, N = 379) = 21.35$, $p < .001$, $\phi_c = .24$. The misinformation effect (increase from 23.1% to 49.3%; $z = 4.64$, $p < .001$, $h = 0.56$) and the warning effect (decrease from 49.3% to 30.7%; $z = 3.12$, $p = .002$, $h = 0.38$) both persisted. The error rate was not significantly higher in the misinformed and warned group (30.7%) than in the control group (23.1%), $z = 1.31$, $p = .190$, $h = 0.17$, which indicates that the misinformation effect was almost eliminated.

The association remained significant for the peripheral item, too, $\chi^2(2, N = 379) = 12.74$, $p = .002$, $\phi_c = .18$. As in the recognition task, the misinformation effect was significant (increase from 11.1% to 29.1%; $z = 3.79$, $p < .001$, $h = 0.46$), whereas the warning had no effect (decrease from 29.1% to 24.6%; $z = 0.82$, $p = .412$, $h = 0.10$). The misinformation effect was also found in the misinformed group that was warned (increase from 11.1% to 24.6%; $z = 2.71$, $p = .007$, $h = 0.36$).

Also with the SM task, the difference of the differences for the misinformation effect and for the warning effect both were not significant, difference of differences = 8.3%, $z = 1.13$, $p = .260$, and difference of differences = 14.1%, $z = 1.74$, $p = .081$, respectively.

Finally, the error rates were compared for the central item and the peripheral item in the misinformed groups. As can be seen from Figure 2, the effect found in the recognition task was reversed in the SM task; the error rate was significantly higher in the central item (41.2%) than in the peripheral item (27.1%), difference = 14.1%, 95% CI [5.4, 22.3], $z = -3.45$, $p = 001$, $h = -0.31$. Post-hoc analyses showed that the effect was significant in the misinformed and unwarned group, but not in the misinformed and warned group, difference =

20.3%, 95% CI [8.4, 31.9], $z = -3.65$, $p < 001$, $h = -0.42$, and difference = 6.1%, 95% CI [-6.9, 17.9], $z = -1.04$, $p = .299$, $h = -0.15$), respectively.

Figure 2

SM Answers in the Warned Group

To gain further insight into why misinformation is reported even in the warned group, false reports in that group were analysed more closely (see Table C1 for an overview). With the central item, false reports were attributed in 85.3% of all cases to the video or the video and the partner, in 12.2% of the cases to the partner only (correct attribution), and in 2.4% of the cases to pure guesswork. Concerning the peripheral item, 36.9% of all erroneous answers were attributed to the video or the video and the partner, 34.2% were correctly attributed to the partner only, and 28.9% to pure guesswork. An overview of the results and the proportion of the different errors can be seen in Figure 3.

Figure 3

Additional Analyses

Response patterns in critical and comparison items in the control group. Working with only two critical items raises the important question whether these two items' response patterns in the memory tests are comparable to those for other central and peripheral items of the original event (comparison items; see Pretest 1 and Table A1). To answer this question, the response patterns for the central and peripheral critical and comparison items were contrasted in the collapsed control conditions. As Figure D1 illustrates, the two critical items' response patterns and their respective comparison items' patterns are quite similar. Furthermore, it becomes clear that the patterns between the central and the peripheral items are fairly divergent.

In general, the central items have a larger proportion of correct recognition answers, and guesswork was rather rare, especially for the central critical item. However, in relation to the comparison items, the central critical item seems to be more susceptible to intrusions: Almost 24% of the participants in the control group indicated that a kick, which was *not* suggested to them in the PEI, had been part of the original event, compared to 11% in the central comparison items. In contrast, the peripheral item elicited a large number of correct and incorrect answers, which can be attributed to pure guesswork. Correct answers were somewhat more frequent in the peripheral critical item than in the comparison items, whereas intrusions occurred comparably often (about 12%). In general, correct recognition answers that were attributed to the incorrect source (e.g. to the partner, although he or she did not actually mention the information) occurred rarely. In sum, the response patterns for the two critical items seem to be quite representative of their respective groups of comparison items.

Overall performance on the comparison items. Participants in the three groups did not differ in the number of correctly answered comparison items (misinformed and unwarned group: M = 16.30, SD = 1.95; misinformed and warned group: M = 16.17, SD = 1.87; control group: M = 16.21, SD = 1.92; $F(2, 426) = 0.21$, p = .811, partial $\eta^2 < .01$).

Discussion

In this research, we sought to better understand the reasons why people report misinformation about central and peripheral aspects of an original event in a recognition task. Therefore, we examined what impact a warning had on the error rate in misinformed participants. We reasoned that a warning reduces false reports due to deliberation or the recency bias by encouraging participants to rely on their own memory and/or to check the source of their memory carefully. A post-recognition source-monitoring (SM) task provided insight into participants' motives for reporting misinformation in the recognition task and was assumed also to be helpful in differentiating between different types of errors.

The Misinformation Effect in the Recognition Task

The misinformation effect (see e.g. Loftus, 2005) was replicated in the recognition task for both items. The error rate in the misinformed and unwarned group was more than twice that in the control group, with an increase of 33% (from 24% to 57%) in the central item and 41% (from 34% to 75%) in the peripheral item. Contrary to other studies (e.g. Dalton & Daneman, 2006), but in agreement with the results reported by Luna and Migueles (2005, 2009), the misinformation effect was not significantly stronger for the peripheral item than for the central item. Of course, it is of interest to examine what the findings on the effect of a warning and on participants' attributions in the SM can tell us about the underlying causes of the misinformation effect.

The Effect of a Warning

We reasoned that, if the misinformation effect observed in the unwarned group (see above) is to some extent due to deliberation and/or a recency bias, the effect should be reduced by a warning. For the central item we found a significant reduction of false reports because of a warning (decrease of the error rate by 19%), although the effect was not completely eliminated. Hence, we conclude that errors other than deliberation or the recency bias seem to be of little importance for the misinformation effect with respect to the central item. But even after a warning some people still reported the central item, that is, they probably developed a false belief. In contrast to the central item, we could not detect a significant effect of the warning for the peripheral item. Hence, deliberation and the recency bias seem to play only a minor role for that item.

One reason why errors due to deliberation and/or the recency bias affect the misinformation effect for the central item so much more strongly than that for the peripheral item may be the stronger memory trace for the central aspect of the original event. In fact, several of our findings lend circumstantial support to the notion that participants had a stronger memory trace for the central item. First, the participants in Pretest 1 reported having paid considerably more attention to the central aspect of the original event (*Mdn* = 5.00) than to the peripheral aspect (*Mdn* = 1.00). Second, more than 55% of the participants in the control group reported that their recognition answer relating to the peripheral item had been pure guesswork, whereas for the central item this applied to only 3% of the cases. Third, although only at a descriptive level, the participants in the control group answered the recognition question concerning the peripheral item incorrectly more often than that concerning the central item. Thus, participants capitalise on the motivation engendered by a warning to strive for a thorough memory search and evaluation and/or to qualify demand characteristics and informational influence, if they have at least some memory of the original event (see also Christiaansen & Ochalek, 1983).

Insights from the Source-Monitoring Task

A very interesting finding was that the false recognition answers in the misinformed groups for the peripheral item were much more often emended in the SM task than those for the central item. Although the results from the recognition task suggest that the peripheral item (error rate of 77%) is more error prone than the central item (56%), this relation is inverted when the SM answers are considered (28% and 48%, respectively).

First of all, this finding allows important conclusions with respect to the underlying causes of false reports of the peripheral misinformation. As mentioned earlier, a weak memory trace may either foster errors of false belief because participants do not detect the discrepancy between their own memory and the PEI and integrate the misinformation into their memory of the original event (see Tousignant et al., 1986), or it may promote best-guess errors as participants have no memory of the original event and knowingly report the information from the PEI. We found that the misinformation effect was not completely eliminated by the SM task. This supports the idea that exposure to the peripheral piece of misinformation evoked false beliefs, at least to some extent. However, the marked decrease of the peripheral item's error rate from over 70% (recognition task) to about 25% (SM task) in the misinformed groups strongly suggests that a considerable proportion of these false reports were due to errors other than false beliefs. The analyses of warned participants' SM answers after they had reported the peripheral misinformation in the recognition task revealed that best-guess errors and pure guesswork were frequent causes of false reports. Hence, best-guess errors, pure guesswork, and false beliefs contribute to a similar extent to the misinformation effect observed in the recognition task (see Figure 3).

Regarding the central item, as mentioned, a strong effect of a warning was observed in the recognition task. Assuming that a warning and the SM task have a similar function, as we outlined in the introduction it follows that they both should reduce errors due to deliberation and the recency bias: While warned participants have already avoided these two errors in the recognition task, unwarned participants are expected to take the chance to emend their incorrect recognition answer by attributing the misinformation to their partner in the SM task. However, we found that the significant misinformation effect observed in the recognition task (56%) with the unwarned participants remained almost stable in the SM task (48%). In terms of our classification, this means that they developed a false belief, which is not in line with the pronounced effect of a warning. This conflicting result is quite surprising and raises the question of whether our warning and SM task are truly functionally equivalent.

Different Functions of the Warning and the Source-Monitoring Task

There are several possible explanations for the different effects of the warning and the SM task in our experiments. On the one hand, the SM answer was assessed within subjects after the recognition answer. Hence, participants may have been reluctant to admit having guessed or relied on their partner, especially concerning such a central aspect of the original event. Yet it may be queried whether these needs or feelings really override the need to answer accurately. Alternatively, participants may actually have developed a false belief whilst elaborating on the recognition question. The results from Experiment 3 (see supplemental online material) suggest that such false beliefs do not evolve generally after answering a recognition question, that is, even if the question was answered correctly. Hence, in order to develop a false belief, participants not only have to elaborate on the question but must also accept the misinformation. This acceptance may be especially likely if participants are not warned, because a warning may help them to promptly reject the misinformation and thereby protects them from developing a false belief.

On the other hand, as Chambers and Zaragoza (2001) argue, the warning may have an additional effect that the SM task does not have. Whilst both measures may have impelled participants to thoroughly monitor the source of information (see also Echterhoff et al., 2005), only the warning may additionally have entailed a more critical consideration of the misinformation, which for instance may have caused a shift in the participants' source decision criterion (cf. source monitoring framework; Johnson, Hashtroudi, & Lindsay, 1993). That is, participants who are warned may set a higher threshold on the quality requirements of their memory (e.g. the memory has to include more visual/auditory details) before attributing it to the original event (see also Chambers & Zaragoza, 2001). In our SM task, participants were not explicitly instructed that some information might have occurred in the

PEI exclusively, which may be an explanation why no such shift occurred. However, this contention is undermined by the fact that Chambers and Zaragoza (2001) used an instruction that clearly stated that some information occurred in the PEI only and also found that an explicit warning in addition to the SM instruction was superior to the SM instruction only.

Limitations and Strengths of the Study

The assessment of the recognition and the SM tasks within participants is a major limitation of our study, in that the exact reasons for the different effects of the warning and the SM task remain unclear. However, our aim of examining the roles of different kinds of errors could only be achieved with a within-subject design. One solution to this problem could be to expand the design by also testing groups that answer only source-monitoring questions (e.g. akin to the SM questions introduced in Experiment 3).

Another limitation is that the two pieces of misinformation were not presented randomly, which might have had an impact on the misinformation effect—especially as the first piece of misinformation blatantly contrasts with the original event (Loftus, 1979). However, if anything, systematically presenting the central misinformation first would have worked against our hypothesis of finding a misinformation effect with a peripheral item. Finally, the SM answers could not be interpreted unambiguously for two reasons. First, some participants may have chosen the *I guessed* option in the SM task because they were unsure of the source from which they remembered the information, and not because their answer in the recognition task had been pure guesswork. Second, participants might occasionally have guessed the answer in the SM task like flipping a coin (cf. source guessing, see e.g. Higham, 1998). In future studies, it would be interesting to add a fifth response option (e.g. *I know it occurred somewhere in the experiment, but I don't know where*, Higham, 1998) to the SM task to filter this type of guessing error.

A consequence of working with a central item is that we could not include many more items, which limits the generalisability of the results. Too many discrepancies between the original event and the PEI could have served as an unintended, implicit warning, as these discrepancies may have raised the participants' suspicions. In general, the additional analyses suggest that the critical items were fairly representative for their respective comparison items. Yet, the central item was rather susceptible to intrusions even in the absence of misinformation. This may have been because the central item was not clearly incompatible with the original event. Although one might argue that remembering that the victim was kicked while lying on the floor is incompatible with remembering that he was not touched after he fell to the ground, participants may have reasoned that they had merely overlooked this very action (see also Chambers & Zaragoza, 2001). Therefore, it remains to be seen whether the results can be generalised to items that are clearly contradictory and/or have a lower general susceptibility to intrusions. Another query over generalisability is the absence of the warning effect in Experiment 3, which was run at the University of Kiel (Germany), whereas the other experiments were run at the University of Bern (Switzerland). In a post hoc analysis, we could not detect significant differences between the samples, for instance concerning the trust in their own memory or in that of their partner, that would explain the inconsistent finding. Thus, it remains unclear whether the inconsistency in the warning effect was due to sampling errors or to actual cultural differences between German and Swiss people, such as peculiarities in the use of language. Last but not least, the results may very well be different if other types of warnings, such as enlightenment (e.g. Oeberst & Blank, 2012) or the logic of opposition (Lindsay, 1990), were used.

Our experiments also have several strengths. Despite the difficulties associated with investigating central, crime-relevant items, the examination of the misinformation effect for such items is crucial, as it increases the ecological validity of the studies. Real-life examples such as the Oklahoma bombing in 1993 or the murder of Anna Lindh in 2003 indicate that witnesses are also susceptible to rather central misinformation. Nonetheless, many studies focus on the effect that peripheral misinformation has on memory reports (see e.g. Luna & Migueles, 2006, cited in Luna & Migueles, 2009). Our results support the view that studying both types of items is vital and further improves our understanding of the mechanisms that underlie false reports. A methodological advantage of our studies is the use of a cover story that does not centre upon memory, because this avoided the possibility that participants watched the video with the explicit expectation of subsequent memory questions. Furthermore, replications were run to gain a clearer picture of the actual effects. This was especially important, as we analysed frequencies, which requires a substantial sample size to achieve a satisfying power and adequate estimates of the effects.

Implications and Future Research

The results clearly show that the reasons for false reports in a recognition task vary broadly depending on the centrality of the items, and they emphasise the importance of taking account of this aspect in future studies. Acquiring knowledge of the extent to which and the reasons why false reports have to be expected after exposure to misinformation is valuable for practical purposes. Depending on the causes of false reports, they can be prevented by certain measures (e.g. a warning or asking participants about the source of their memory). In the warning condition of our experiments, participants were merely informed about the possibility of having been exposed to incorrect information, which nonetheless strongly reduced the occurrence of false reports for the central item. In principle, this could also be done in a police interrogation. But caution should be exercised with such action, as the reporting of correct information may be collaterally damaged (tainted truth effect; Echterhoff et al., 2007). Hence, additional research is needed to establish the potential detrimental effects that a warning can have before, for example, scripting an information leaflet about the misinformation effect in legal practice.

Moreover, the approach presented here may be a promising starting point for further research which may also help to explain other findings in the field on the misinformation effect. For instance, the fact that the warning only reduced errors in the central item may be an explanation for the inconsistent findings in warning studies (see e.g. Blank & Launay, 2014; Echterhoff et al., 2005): If the critical items used in a study are too peripheral or if participants did not pay enough attention to the original event for whatsoever reason, it may simply be not possible for them to benefit from a warning.

It is important to test the generalisability of the results reported here to other instances of misinformation and to examine the role other characteristics of misinformation may play. But, as mentioned, the number of central items that can be researched within participants is limited, and it is challenging to find items that vary only with respect to their centrality but not in other characteristics, such as schema consistency or the general susceptibility to intrusions. One approach to circumventing this problem is to experimentally manipulate the attention participants can or do pay towards certain aspects of the original event, either directly by a divided attention task, or indirectly by using specific instructions or cover stories to direct their attention.

Another important question is why the warning and the SM task did not have comparable effects. To decide whether participants actually develop a false belief or whether the warning causes a shift in the response threshold, some participants who are not warned before answering the recognition questions could be provided with an explicit warning before the SM task in a future study. If this warning does not have the same effect as a warning given before the recognition question, this would indicate that the elaboration on the recognition question may be accompanied by some processes that foster the development of false beliefs.

Conclusion

Examining the impact that misinformation on central aspects of an event has on people's memory and memory reports is practically relevant, as various cases of witnesses' misled testimonies illustrate. As our results emphasise, it is crucial to consider items of varying centrality, because the reasons for false reports differ depending on whether the misinformation concerns central or peripheral aspects of the original event. For instance, best guesses and false beliefs seem to be especially important for the peripheral item, while deliberation and/or the recency bias are presumably vital for the central item, although the inconsistent effects of the warning and the SM task impede a final conclusion with respect to this latter item. The results nevertheless show that the combined investigation of the effects of a warning and a post-recognition SM task is a promising and worthwhile approach to directly investigating the role of different errors in the reporting of misinformation. The combined investigation especially improves the identification of false beliefs, as it helps to filter best-guess errors, which would not be possible with a warning alone. Last but not least, these insights improve both the assessment of the impact that misinformation has on witnesses' testimonies and our understanding of how to prevent or identify the reporting of misinformation.

References

Allan, K., Midjord, J. P., Martin, D., & Gabbert, F. (2012). Memory conformity and the perceived accuracy of self versus other. *Memory & Cognition, 40,* 280–286. doi:10.3758/s13421-011-0141-9

Belli, R. F. (1989). Influence of misleading postevent information: Misinformation interference and acceptance. *Journal of Experimental Psychology: General, 118,* 72–85. doi:10.1037/0096-3445.118.1.72

Belli, R. F., Lindsay, D. S., Gales, M. S., & McCarthy, T. T. (1994). Memory impairment and source misattribution in postevent misinformation experiments with short retention intervals. *Memory & Cognition, 22,* 40–54. doi:10.3758/BF03202760

Belli, R. F., & Loftus, E. F. (1996). The pliability of autobiographical memory: Misinformation and the false memory problem. In D. C. Rubin (Ed.), *Remembering our past* (pp. 157–179). New York: Cambridge University Press

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. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 356–367. doi:10.1037/0278-7393.18.2.356

Blank, H. (1998). Memory states and memory tasks: An integrative framework for eyewitness memory and suggestibility. *Memory, 6,* 481–529. doi:10.1080/741943086

Blank, H. (2009). Remembering: A theoretical interface between memory and social psychology. *Social Psychology, 40*, 164–175. doi:10.1027/1864-9335.40.3.164

Blank, H., & Launay, C. (2014). How to protect eyewitness memory against the misinformation effect: A meta-analysis of post-warning studies. *Journal of Applied Research in Memory and Cognition, 3*, 77–88. doi:10.1016/j.jarmac.2014.03.005

Bodner, G E., Musch, E., & Azad, T. (2009). Reevaluating the potency of the memory conformity effect. *Memory & Cognition, 37*, 1069–1076. doi:10.3758/MC.37.8.1069

Bortz, J., Lienert, G. A., & Boehnke, K. (2008). *Verteilungsfreie Methoden in der Biostatistik* [Non-parametric methods in biostatistics] (2nd rev. ed.). Berlin, Germany: Springer.

Chambers, K. L., & Zaragoza, M. S. (2001). Intended and unintended effects of explicit warnings on eyewitness suggestibility: Evidence from source identification tests. *Memory & Cognition, 29*, 1120–1129. doi:10.3758/BF03206381

Christiaansen, R. E., & Ochalek, K. (1983). Editing misleading information from memory: Evidence for the coexistence of original and postevent information. *Memory & Cognition, 11*, 467–475. doi:10.3758/BF03196983

Christianson, S. A. (1992). Emotional stress and eyewitness memory: A critical review. *Psychological Bulletin, 112*, 284–309. doi:10.1037/0033-2909.112.2.284

Chrobak, Q. M., & Zaragoza, M. S. (2013). The misinformation effect: Past research and recent advances. In A. M. Ridley, F. Gabbert, & D. J. La Rooy (Eds.), *Suggestibility in legal contexts: Psychological research and forensic implications* (pp. 21–44). Chicester: Wiley-Blackwell.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* New York, NY: Academic Press.

Dalton, A. L., & Daneman, M. (2006). Social suggestibility to central and peripheral misinformation. *Memory, 14,* 486–501. doi:10.1080/09658210500495073

Deutsch, M., & Gerard, H. G. (1955). A study of normative and informational social influence upon individual judgement. *Journal of Abnormal and Social Psychology, 59*, 204-209. doi:10.1037/h0046408

Echterhoff, G., Groll, S., & Hirst, W. (2007). Tainted truth: Overcorrection for misinformation influence on eyewitness memory. *Social Cognition, 25*, 367–409. doi:10.1521/soco.2007.25.3.367

Echterhoff, G., Hirst, W., & Hussy, W. (2005). How eyewitnesses resist misinformation: Social postwarnings and the monitoring of memory characteristics. *Memory & Cognition, 33*, 770–782. doi:10.3758/BF03193073

Frost, P., & Weaver, C. A. (1997). Overcoming misinformation effects in eyewitness memory: Effects of encoding time and event cues. *Memory, 5*, 725–740. doi:10.1080/741941545

Gabbert, F., & Hope, L. (2013). Suggestibility and memory conformity. In A. M. Ridley, F. Gabbert, & D. J. La Rooy (Eds.), *Suggestibility in legal contexts: Psychological research and forensic implications* (pp. 63–83). Chicester: Wiley-Blackwell.

Heath, W. P., & Erickson, J. R. (1998). Memory for central and peripheral actions and props after varied post-event presentation. *Legal and Criminological Psychology, 3*, 321–346. doi:10.1111/j.2044-8333.1998.tb00369.x

Higham, P. A. (1998). Believing details known to have been suggested. *British Journal of Psychology, 89*, 265–283. doi:10.1111/j.2044-8295.1998.tb02684.x

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. doi:10.1016/j.actpsy.2014.06.006

Horry, R., Palmer, M. A., Sexton, M. L., & Brewer, N. (2012). Memory conformity for confidently recognized items: The power of social influence on memory reports. *Journal of Experimental Social Psychology, 48*, 783–786. doi:10.1016/j.jesp.2011.12.010

Ibabe, I., & Sporer, S. L. (2004). How you ask is what you get: On the influence of question form on accuracy and confidence. *Applied Cognitive Psychology, 18,* 1–16. doi:10.1002/acp.1025

Johnson, M. K., Hashtroudi, S., & Lindsay, D. (1993). Source monitoring. *Psychological Bulletin, 114,* 3–28. doi:10.1037/0033-2909.114.1.3

Lane, S. M. (2006). Dividing attention during a witnessed event increases eyewitness suggestibility. *Applied Cognitive Psychology, 20*, 199–212. doi:10.1002/acp.1177

Lindsay, D. S. (1990). Misleading suggestions can impair eyewitnesses' ability to remember event details. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 1077–1083. doi:10.1037/0278-7393.16.6.1077

Lindsay, D. S., & Johnson, M. K. (1989). The eyewitness suggestibility effect and memory for source. *Memory & Cognition, 17,* 349–358. doi:10.3758/BF03198473

Loftus, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology, 7,* 560–572. doi:10.1016/0010-0285(75)90023-7

Loftus, E. F. (1979). Reactions to blatantly contradictory information. *Memory & Cognition, 7*, 368–374. doi:10.3758/BF03196941

Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory, 12*, 361–366. doi: 10.1101/lm.94705

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4*, 19–31. doi:10.1037/0278-7393.4.1.19

Luna, K., & Migueles, M. (2005). Effecto del tipo de información sugerida en el paradigma de la información postevento. [Effect of the types of suggestive information in the misinformation paradigm.] *Revista de Psicología General y Aplicada, 58*, 309–321.

Luna, K., & Migueles, M. (2009). Acceptance and confidence of central and peripheral misinformation. *The Spanish Journal of Psychology*, *12*, 405–413. doi:10.1017/S1138741600001797

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. doi:10.1037/0096-3445.114.1.1

Oeberst, A., & Blank, H. (2012). Undoing suggestive influence on memory: The reversibility of the eyewitness misinformation effect. *Cognition, 125*, 141–159. doi:10.1016/j.cognition.2012.07.009

Roebers, C. M., & McConkey, K. M. (2003). Mental reinstatement of the misinformation context and the misinformation effect in children and adults. *Applied Cognitive Psychology, 17,* 477–493. doi:10.1002/acp.886

Roos af Hjelmsäter, E., Granhag, P. A., Strömwall, L. A., & Memon, A. (2008). The effects of social influence on children’s memory reports: The omission and commission error asymmetry. *Scandinavian Journal of Psychology, 49*, 507–513. doi:10.1111/j.1467-9450.2008.00679.x

Shapiro, L. R., Blackford, C., & Chen, C.-F. (2005). Eyewitness memory for a simulated misdemeanour crime: The role of age and temperament in suggestibility. *Applied Cognitive Psychology, 19,* 267–289. doi:10.1002/acp.1089

Tousignant, J. P., Hall, D., & Loftus, E. F. (1986). Discrepancy detection and vulnerability to misleading postevent information. *Memory & Cognition, 14*, 329–338. doi:10.3758/BF03202511

Wright, D. B., & Stroud, J. N. (1998). Memory quality and misinformation for peripheral and central objects. *Legal and Criminological Psychology, 3*, 271–286. doi:10.1111/j.2044-8333.1998.tb00366.x

Zaragoza, M. S., Belli, R. S., & Payment, K. E. (2007). Misinformation effects and the suggestibility of eyewitness memory. In M. Garry & H. Hayne (Eds.). *Do justice and let the sky fall: Elizabeth F. Loftus and her contributions to science, law, and academic freedom* (pp. 35–63). Hillsdale, NJ: Lawrence Erlbaum Associates.

Zaragoza, M. S., & Koshmider, J. W. III (1989). Misled subjects may know more than their performance implies. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 246–255. doi:10.1037/0278-7393.15.2.246

Zaragoza, M. S., & Lane, S. M. (1994). Source misattributions and suggestibility of eyewitness memory. *Journal of Experimental Psychology: Learning, Memory and Cognition 20*, 934–945. doi:10.1037/0278-7393.20.4.934

Appendix A

Table A1

Centrality Index of the Critical and Comparison Aspects of the Original Event

Central Aspects	Centrality Index (*Mdn*)
Number of witnesses at the scene	6.50
Body height of the victim relative to the perpetrators	5.50
Sex of the witnesses	5.50
Threat the victim made against the perpetrators (will call the police)	5.50
How victim was treated when lying on the ground	**5.00**
Objects taken from the victim	5.00
Language spoken by the victim	4.50
Body area in which the victim was punched	4.00
Place where the passive perpetrator was standing during the incident	4.00
Location of the pocket in which the victim stored his wallet	3.00
Type of automat/booth at the location	3.00
Peripheral Aspects	Centrality Index (*Mdn*)
Number of perpetrators with headwear	2.50
Headwear worn by the passive perpetrator	**2.00**
Number of witnesses carrying a bag	2.00

Pocket in which perpetrator 1 put the banknote he allegedly wanted to change	2.00
Type of facial hair worn by passive perpetrator	2.00
Colour of perpetrator 1's coat	1.50
Content of window display	1.50
Head covering of the witness who knelt next to the victim	1.50
Number of perpetrators wearing jeans	1.50
Number of perpetrators with a scarf	1.50
Type of shoes worn by perpetrator 1	1.50

Note. Overview on the content and centrality index ratings from Pretest 1 (scale ranging from 1 to 7) of the central and peripheral aspects on which participants were asked questions at the memory test (translated from the German original version). The two critical aspects are printed in bold.

Appendix B

Table B1

Error Rates in the Recognition Task and Remaining Errors in the Source-Monitoring Task in the Three Conditions for the Central and the Peripheral Item per Experiment and Overall

	Misinformed and unwarned group				**Misinformed and warned group**				**Control group**			
	Exp. 1	Exp. 2	Exp. 3	**Overall**	Exp. 1	Exp. 2	Exp. 3	**Overall**	Exp. 1	Exp. 2	Exp. 3	**Overall**
Central item												
Recognition task	59.4	53.4	57.7	**56.9**	27.9	29.4	56.8	**37.7**	36.6	18.0	15.4	**23.9**
SM task	48.4	48.3	53.8	**49.3**	25.6	25.5	55.0	**30.7**	36.6	18.0	11.5	**23.1**
Peripheral Item												
Recognition Task	68.8	82.8	73.1	**74.7**	62.8	74.5	68.2	**68.8**	41.5	28.0	34.6	**34.2**
SM task	28.1	27.6	34.6	**29.1**	30.2	25.5	10.0	**24.6**	17.1	4.0	15.4	**11.1**

Note. All figures represent percentages. Experiment 1 also includes the data from the posttest.

Appendix C

Table C1

Source Attributions of Misinformation Selected in Response to the Recognition Question

	Central item			**Peripheral item**		
	Misinformed and unwarned group	Misinformed and warned group	Control group	Misinformed and unwarned group	Misinformed and warned group	Control group
No of cases[1]	86	41	28	112	76	40
Video only	26.7	39.0	67.9	21.4	21.1	32.5
Video and partner	58.1	46.3	28.6	17.0	15.8	0.0
Guess	3.5	2.4	0.0	26.8	28.9	67.5
Partner only[2]	11.6	12.2	3.6	34.8	34.2	0.0

Note. All figures represent percentages, except for the number of cases. The results refer to the overall analyses (N = 379). [1]Number of participants who chose the misinformation response in the recognition task. [2]Corresponds to a correct source attribution in the misinformed groups.

Appendix D

Figure D1. Response patterns for the two critical items and their respective comparison items in the collapsed control groups. Recognition correct and SM incorrect: The correct answer is attributed to the partner (PEI), although the partner did not mention this information.

Table 1

Theoretical Assumptions Concerning the Effect of a Warning and the Attribution of Recognition Errors for the Different Types of Errors

Assumed type of error	Effect of warning	Correction in the SM task
Deliberation	Yes	Yes
Recency bias	Yes	Yes
Best guess	No	Yes
False belief	No	No

Note. For example, errors due to deliberation are expected to be reduced by a warning. Furthermore, if such a report occurs (e.g. because participants were not warned), the incorrect recognition answer should be attributed to the partner and not to the original event by definition. In this sense, participants correct their erroneous recognition answer. An additional situation not included in the table is the case of pure guesswork, in which the misinformation is reported but participants indicate in the SM task that they guessed.

Table 2

Comparing Overall Error Rates in the Recognition Task and Remaining Errors in the SM Task for the Central and the Peripheral Misinformation Item in the Three Groups.

	Central item			Peripheral item		
	Effect	Diff	95% CI	Effect	Diff	95% CI
Recognition task						
ME	Yes	33.0	[21.0, 45.0]	Yes	40.5	[27.7, 52.1]
Effect of warning	Yes	19.2	[8.2, 30.5]	No	5.9	[-4.7, 16.8]
Elimination of ME	No	13.7	[2.3, 25.6]	No	34.7	[21.4, 47.2]
SM task						
ME	Yes	26.2	[12.5, 38.6]	Yes	18.0	[7.0, 28.0]
Effect of warning	Yes	18.6	[3.1, 30.9]	No	4.5	[-7.5, 15.3]
Elimination of ME	Yes	7.6	[-5.0, 19.1]	No	13.5	[2.7, 24.7]

Note. All figures represent percentages and refer to the overall analyses. *Diff* = difference between the two respective groups; CI = confidence interval. *Effect* refers to whether the data is supportive of the research question. *ME* (misinformation effect): Significant difference between misinformed and unwarned vs. control group. *Effect of warning*: Significant reduction of the error rate by a warning (difference between misinformed and unwarned vs. misinformed and warned group). *Elimination of ME*: Elimination of the misinformation effect because of a warning (difference between misinformed and warned group vs. control group is insignificant).

Figure 1. Overview of the experimental approach. *The warning was manipulated between subjects in the misinformed groups. **A control group was exposed neither to misinformation nor to a warning.

Figure 2. Overall error rates in the recognition task and errors remaining in the SM task (incorrect answer in the recognition task and attribution to the original event) in the different groups.

Figure 3. Error rate in the misinformed and unwarned group explained by different types of errors for the central and the peripheral item. False beliefs are the errors that persist even in the SM task, thas is, the misinformation is attributed to the video or the video and the partner. R = Recognition task, SM = SM task.

Online Supplementary Materials

Experiment 1

Method

Participants. Overall, 133 participants were tested. Of these, 18 participants had to be excluded due to not remembering the warning (8), problems with the video presentation (3), or not believing at all that the information was actually generated by a former participant (7). The final groups consisted of 47 (misinformed and unwarned), 43 (misinformed and warned), and 25 (control) participants. Mean age was M = 22.36 years (SD = 7.31), and 86 participants were women. Sixty-seven participants were undergraduate psychology students; the others were students from other disciplines.

Design. The experiment used a mixed design with a three-level between factor (*misinformed and unwarned, misinformed and warned, control*), and a two-level within factor (*central* and *peripheral*).[8] In the control condition participants were neither warned nor exposed to misleading information. The dependent variables were frequencies (error rates in the recognition and SM task).

Results

Comparisons of proportions, Cohen's *h* (with small, medium, and large effects at 0.20, 0.50, and 0.80, respectively; Cohen, 1988), and Cramér's phi (with small, medium and large effects at .10, .30, and .50, respectively; Cohen, 1988) were calculated by hand. The

[8] In Experiments 1 and 2, we also manipulated the credibility of the partner with the aim of investigating the effects of a warning before the exposure to misinformation (prewarning) in combination with those of a warning after the misinformation (postwarning). Participants were informed that their partner had either seen the video twice (high credibility) or had only read another participant's memory report of the incident (low credibility). However, the manipulation checks revealed that the manipulation did not work as expected; participants rated their partner as being equally knowledgeable as they were in the low-credibility condition, although we had excluded participants if they either did not remember the credibility manipulation correctly or did not believe that the information came from a previous participant (see French, Garry, & Mori, 2011, who found that relative, not absolute, judgements of credibility influence how susceptible participants are to misinformation). Therefore, it was unsurprising that credibility did not significantly impact the error rates. As the manipulation check suggests that the manipulation itself did not work, the credibility conditions were collapsed for all statistical analyses (more information on the results is available on request; please contact the first author). In Experiment 3, only the high-credibility condition was realised.

statistical software R was used to calculate exact tests, which are reported additionally if the conditions for *z*-Tests are not fully met, as well as 95% confidence intervals for the differences in proportions (R package *ExactCIdiff*). Please note that there may be rounding differences, especially for the differences calculated with the R package. All other statistical analyses were conducted with SPSS 21.0. In all three experiments, statistical tests are reported two-tailed, and, because of the multiple testing per research question, a Bonferroni correction was applied, resulting in an alpha level of .017. Furthermore, the alpha level was adjusted for additional analyses individually in each experiment.

In Experiment 1, data from 115 participants were included. Table B1 (see main article) gives an overview of the error rates in the recognition task and of the errors that persisted in the SM task for the two items. The adjusted alpha level for additional analyses was .025.

Baseline error. In the recognition task, the baseline error (control group) did not significantly differ between the central item (52.0%) and the peripheral item (48.0%), difference = 4.0%, 95% CI [-25.1, 32.5], z = -0.28, p = .781, h = -0.08, binomial test p = 1.000. In the SM task, in contrast, the error rate slightly tended to be higher in the central item (52.0%) than in the peripheral item (20.0%), difference = 32.0%, 95% CI [2.4, 57.1], z = -2.14, p = .033, h = -0.61, binomial test p = .057.

Recognition task. The chi-square test of independence conducted on the central item revealed a significant association between the three groups and the accuracy of the answer (correct/incorrect), $\chi^2(2, N = 115) = 8.50$, p = .014, ϕ_c = .27. Thus, column proportions of the different groups were compared (alpha level = .017). No significant misinformation effect was observed, as the comparison of the error rate in the control group (52.0%) and the misinformed and unwarned group (57.5%) was not significant, difference = 5.4%, 95% CI [-

18.9, 30.6], z = 0.44, p = .658, h = 0.11. Nonetheless, the warning significantly reduced the error rate in the misinformed group from 57.4% to 27.9%, difference = 29.5%, 95% CI [8.5, 48.9], z = 2.97, p = .003, h = 0.61. Although the difference between the control group (52.0%) and the misinformed and warned group (27.9%) failed to reach significance, the error rate was lower in the warned condition than in the control condition at a descriptive level, difference = 24.1%, 95% CI [-1.0, 47.1], z = -1.99, p = .047, h = -0.50.

The test of independence for the peripheral item was not significant, $\chi^2(2, N = 115)$ = 1.93, p = .382, ϕ_c = .13, i.e., no statistically significant differences were observed in the error rates (control: 48.0%; misinformed and warned: 62.8%; misinformed and unwarned: 63.8%).

Finally, the error rates in the central and the peripheral item in the misinformed groups were analysed. Participants committed significantly fewer errors with the central item (43.3%) than with the peripheral item (63.3%), difference = 20.0%, 95% CI [3.8, 34.9], z = 2.65, p = .008, h = 0.40.

SM task. The chi-square test of independence for both the central item and the peripheral item, failed to reach significance, $\chi^2(2, N = 115)$ = 5.66, p = .059, ϕ_c = .22, and $\chi^2(2, N = 115)$ = 1.12, p = .602, ϕ_c = .09, respectively (for the exact error rates in the different groups, see Table B1). Although not significant, the error rate in the central item was as much as 19.1% higher in the misinformed and unwarned than in the misinformed and warned group.

The error rate in the central item (35.6%) and the peripheral item (26.7%) exhibited no significant difference in the collapsed misinformed groups, difference = 8.9%, 95% CI [-6.1, 22.9], z = -1.26, p = .206, h = -0.20. Yet, it should be noted that the effect found with the recognition task was reversed at the descriptive level.

Discussion

A significant misinformation effect was found for neither the central item nor the peripheral item in the recognition task. However, we could not exclude the possibility that a sampling error caused these results. This assumption was reinforced by the fact that the error rate in the central item was 24% higher in the control group than in the misinformed and warned group, although the difference narrowly failed to reach significance. Therefore, the impact of the warning on the error rate in the misinformed and unwarned group was also investigated. The warning was indeed found to reduce the error rate to half of its original size. No significant association between the three groups and the error rate was found with respect to the peripheral item. Thus, the results suggest that the warning reduced the error rate in the recognition task, but only for the central item. In other words, participants seem to have falsely reported the central item at least partially because of deliberation and/or the recency bias, whereas no such evidence was found for the peripheral item.

No misinformation effect was observed for the SM task either. Furthermore, the effect of the warning found for the central item in the recognition task disappeared at a statistical level. Yet, the error rate still was as much as 19% higher in the misinformed and unwarned group than in the warned group. This finding may provide evidence that a warning and an SM task may—at least to some extent—affect different causes of false reports. But before examining this finding more closely, we wanted to test whether it was replicable.

The frequency of false reports in the misinformed groups tended to be higher in the peripheral item than in the central item (difference of 20%). In the SM task, in contrast, the participants committed *more* errors in the central item, although only at a descriptive level. Hence, participants may have benefitted more from the opportunity to revise an incorrect answer in the peripheral item than in the central item.

These first results indicate that a substantial amount of false reports about the central critical aspect of the original event may be due to deliberation and/or the recency bias. For the peripheral item, a quite different picture emerged. As the warning had no error-reducing effect, errors due to deliberation and/or the recency bias probably do not underlie the false reports. Rather, the fact that participants—although only descriptively—benefit more from the SM task for the peripheral item than the central item suggests that a large proportion of the false reports about the peripheral item are due to best guesses or pure guesswork.

The high baseline error of the central item in both the recognition task and the SM task suggests that this item might be quite susceptible to the development of false beliefs even in the absence of misinformation. Nearly half of the participants in the control condition believed that the victim had been kicked, although this was suggested neither in the video nor in the PEI. There are different possible explanations for this finding. First, as the item is rather schema-consistent, it may have been due to some kind of intrusion by activating the schema or script of a typical robbery. This activation may have resulted in a similar effect as in the Deese-Roediger-McDermott (DRM) paradigm (Roediger & McDermott, 1995; see also Roediger, Watson, McDermott, & Gallo, 2001). Second, the forced-choice recognition question could have suggested the presence of a kick to the participants, and participants readily developed a firm belief that this had actually happened. Third, the result may have been due to a sampling error. This last assumption receives some support from the descriptively lower error rate in the misinformed and warned group than in the control group. We also collected more data by running a posttest (N = 33 with a misinformed and unwarned group and a control group) for Experiment 1. The error rate was 12.5% in the control group (N = 16) and 64.7% in the misinformed and unwarned group (N = 17). Hence, we could not replicate the high error rate that was observed in the control group in Experiment 1. However, having two contradictory results does not yet provide information on which of the two findings might have been due to a sampling error.

Experiment 2 was designed to shed light on the causes of the high baseline error rate in the control condition as well as on the robustness of the findings from Experiment 1. To differentiate between the possible explanations for the high baseline error rate, additional open-ended questions were introduced prior to the recognition task. We reasoned that intrusions that occurred during the participants watched the video should be reported in open-ended questions. By contrast, if the high baseline error were due to a possible suggestive effect of the recognition question, the error rate should be rather low in the open-ended questions but high in the recognition task. If, however, the high baseline error was merely a sampling error, it should not be replicated. As it cannot be excluded that the open-ended questions affect the possible suggestive effect of the recognition question, a second control group identical to that in Experiment 1 (no open-ended questions) was included.

Experiment 2

Method

Participants. One hundred and eighty-six students participated in Experiment 2. Overall, 27 participants had to be excluded due to not remembering the warning (12), problems with the video presentation (2), disturbances during the video presentation (2), not believing at all that the information was actually generated by a former participant (4), or knowing the video already (7). The final groups consisted of 58 (misinformed and unwarned), 51 (misinformed and warned), 26 (control with open-ended questions), and 24 (control without open-ended questions) participants. Mean age was M = 22.84 years (SD = 3.41), 129 participants were women, and 34 participants were undergraduate psychology students. The other participants were students from other disciplines.

Design. The design of Experiment 2 was identical to that of Experiment 1, except for the second control group without open-ended questions, which was added as a replication of the control group from Experiment 1. In addition to the recognition and SM task, participants' answers to the open-ended questions were analysed for the control groups (coded as critical item clearly mentioned: *yes* or *no*).

Material and procedure. The material was identical to that of Experiment 1, except for three (two critical, one distractor) additional open-ended questions. The questions allowed participants to report the misinformation but were not suggestive. The question targeting the central item was: *What kind of violence was exerted on the victim during the incident? Please describe the individual actions in more detail.* This question is not considered to explicitly suggest the central misinformation (kick), because several other violent actions, such as being pushed against a wall, were exerted on the victim. The question targeting the peripheral item was: *One of the perpetrators was not directly involved in the attack. Please describe his appearance and his clothing in detail.* Two raters coded all answers as to whether or not the critical items (i.e., the kick in the stomach and/or the pulled-up hood) were clearly mentioned. Inter-rater agreement was 95% for both questions. Discrepancies between raters were discussed and resolved.

The procedure was the same as in Experiment 1, except for the three additional open-ended questions that were answered by all participants except for those in the second control group prior to the recognition task. In the warning condition, the warning was provided before the open-ended questions.

Results

Data from 159 participants were included in statistical analyses (see Table B1 for an overview on the error rates in the different groups). The adjusted alpha level for additional analyses was .008.

Baseline error. A first question of interest was whether the high baseline errors found in Experiment 1 were replicated. First, we tested whether the baseline errors were significantly different for the participants in the control groups with versus without the open-ended questions. As no significant difference was observed (central item: 19.2% and 16.6%; difference = 2.6%, 95% CI [-20.6, 25.0], $z = 0.24$, $p = .813$, $h = 0.07$, Fisher's Exact Test (FET) $p = 1.000$; peripheral item: 26.9% and 29.1%; difference = 2.2%, 95% CI [-22.4, 27.9], $z = 0.18$, $p = .860$, $h = 0.05$, FET $p = 1.000$), the data from the two control groups were collapsed for further analyses. Next, the error rates in the recognition task were compared for Experiment 1 and 2. In Experiment 2, the error rate was significantly lower for the central item (Experiment 1: 52.0%, Experiment 2: 18.0%; difference = 34.0%, 95% CI [11.2, 54.9], $z = 2.99$, $p = .003$, $h = 0.73$), whereas the difference failed to reach significance for the peripheral item (Experiment 1: 48.0%; Experiment 2: 28.0%; difference = 20.0%, 95% CI [3.2, 42.9], $z = 1.69$, $p = .091$, $h = 0.42$). As in Experiment 1, the central item (18.0%) and the peripheral item (28.0%) did not differ significantly in their error rate for the recognition task, difference = 10.0%, 95% CI [-5.7, 25.9], $z = 1.29$, $p = .197$, $h = 0.34$, FET $p = .302$. For the SM task, participants in the control group slightly tended to commit more errors in the central item (18.0%) than the peripheral item (4.0%), difference = 14.0%, 95% CI [2.2, 27.1], $z = -2.30$, $p = .022$, $h = -0.89$, binomial test $p = .039$, although the difference clearly failed to reach significance with the exact test.

Recognition task. As in Experiment 1, the three groups and the accuracy of the answers to the central item were associated significantly, $\chi^2(2, N = 159) = 15.80$, $p < .001$, $\phi_c = .32$. Individual comparisons revealed a misinformation effect, as the error rate increased significantly from 18.0% in the control group to 53.4% in the misinformed and unwarned group, difference = 35.4%, 95% CI [11.0, 52.1], $z = 4.17$, $p < .001$, $h = 0.76$. The warning not only reduced the error rate from 53.4% to 29.4%, difference = 24.0%, 95% CI [4.0, 45.6], $z = 2.63$, $p = .009$, $h = 0.49$; the misinformation effect was even eliminated statistically (misinformed and warned: 29.4%, control: 18.0%; $z = 1.36$, $p = .173$, $h = 0.27$), although a difference of 11.4%, 95% CI [-8.7, 30.6], remained.

For the peripheral item too, a significant association was found between the groups and the commission of an error, $\chi^2(2, N = 159) = 38.84$, $p < .001$, $\phi_c = .49$. There was a significant misinformation effect (increase from 28.0% to 82.8%; difference = 54.8%, 95% CI [38.0, 70.1], $z = 6.80$, $p < .001$, $h = 1.17$). In contrast to the central item, the warning did not significantly reduce the error rate for the peripheral item (decrease from 82.7% to 74.5%; difference = 8.2%, 95% CI [-7.9, 25.8], $z = 1.05$, $p = .294$, $h = 0.20$), and the error rate remained significantly higher in the misinformed and warned group (74.5%) than in the control group (28.0%), difference = 46.5%, 95% CI [27.4, 62.7], $z = 5.28$, $p < .001$, $h = 0.97$.

Over the two misinformed groups, participants committed significantly fewer errors with the central item (42.2%) than with the peripheral item (78.9%), difference = 36.7%, 95% CI [24.0, 48.2], $z = 5.35$, $p < .001$, $h = 0.80$.

SM task. Here too, a significant association was found between the groups and false reports of the central misinformation, $\chi^2(2, N = 159) = 12.65$, $p = .002$, $\phi_c = .28$. The misinformation effect (increase from 18.0% to 48.3%) and the effect of the warning (decrease from 48.3% to 25.5%) were both significant, difference = 30.3%, 95% CI [7.3, 46.8], $z =$

3.55, $p < .001$, $h = 0.66$, and difference = 22.8%, 95% CI [2.8, 43.9], $z = 2.54$, $p = .011$, $h = 0.48$, respectively. As in the recognition task, no significant difference between the misinformed and warned group (25.5%) and the control group (18.0%) was detected, difference = 7.5%, 95% CI [-12.2, 23.9], $z = 0.92$, $p = .359$, $h = 0.18$.

The significant association also persisted in the peripheral item, $\chi^2(2, N = 159) = 11.24$, $p = .004$, $\phi_c = .27$. A misinformation effect remained (increase from 4.0% to 27.6%; difference = 23.6%, 95% CI [8.5, 37.3], $z = 3.63$, $p < .001$, $h = 0.70$, FET $p = .001$), the warning had no significant impact on the error rate (decrease from 27.6% to 25.5%; difference = 2.1%, 95% CI [-15.9, 20.0], $z = 0.25$, $p = .804$, $h = 0.05$), and the misinformation effect was also found in the misinformed and warned group (increase from 4.0% to 25.5%; difference = 21.5%, 95% CI [8.0, 36.2], $z = 3.21$, $p = .001$, $h = 0.66$, FET $p = .004$).

Other than in the recognition task, no significant difference was observed between the error rate in the central (37.6%) and the peripheral item (26.6%), difference = 11.0%, 95% CI [-2.0, 23.4], $z = -1.73$, $p = .083$, $h = -0.25$. although the effect was reversed at the descriptive level, as was the case in Experiment 1.

Open-ended questions. The open-ended questions were introduced to examine the occurrence of intrusions in the control group. A kick (suggested central item) was mentioned by two (7.7%) of the control-group participants, and one person (3.8%) was unsure whether the victim was kicked. No participant spontaneously mentioned the pulled-up hood (suggested peripheral item).

Discussion

In contrast to Experiment 1, a distinct misinformation effect was found in Experiment 2 for both items in the recognition task. Concerning the central item, the main reason for this finding was the lower baseline error in the control group. For the peripheral item, the combination of a lower baseline error and a higher error rate in the misinformed group without a warning was responsible for the effect. As in Experiment 1, a significant warning effect was found for the central item but not for the peripheral item, and the misinformation effect was even eliminated; the difference in the error rate between the misinformed group with a warning and the control group was not significant. By contrast, the misinformation effect in the peripheral item was observed in both misinformed groups.

Both the misinformation effect and the warning effect for the central item persisted in the SM task. The finding that the error rate was nearly 23% higher in the misinformed and unwarned group than in the misinformed and warned group suggests that a warning may reduce some causes of false reports that are not captured by the SM task. If the error-reducing effect of the warning were caused by the same reasons as those of the SM task, the warning effect should not persist in the unwarned group when the error rate observed in the SM task is considered. One possible explanation is that answering the recognition questions activates the original event and the misinformation at the same time. This joint activation may foster false beliefs at a later stage (Zaragoza & Lane, 1994), possibly independently even of whether the initial answer was correct or not (see e.g., Ackil & Zaragoza, 1998, who found that participants develop false memories even for forced fabrications). This possibility was explored in Experiment 3 (see below).

In the recognition task, the error rate in the misinformed groups was 37% higher in the peripheral item than in the central item. Again, the effect was reversed at a descriptive level in the SM task. Hence, as in Experiment 1, the chance to revise an incorrect answer seems to be used more for the peripheral item than for the central item. Overall, the results of Experiment 2 support those of Experiment 1; the causes of false reports seem to be quite different for the central item and the peripheral item. According to our findings, the former seems to be mainly, if not entirely, due to deliberation or the recency bias, whereas in the latter best-guess errors and pure guesswork seem to have a substantial share.

The results of Experiment 2 also give some indication of what may have caused the high baseline error in Experiment 1. As the high error rate was replicated in neither of the two control groups and as the error rate was significantly lower in Experiment 2 than in Experiment 1, a sampling error seems to be a plausible explanation. The results from the posttest additionally support this conclusion. Furthermore, although some spontaneous intrusions, meaning a development of false beliefs in the absence of misinformation, occurred in the answers to the open-ended questions, they were rather rare.

The aims of Experiment 3 were twofold. First, we wanted to know whether main results of the first two experiments could be replicated with a sample from another university in another country. Second, we sought a better understanding of why the significant effect of a warning persisted in the SM task for the central item. To test the idea that the mere act of elaborating on the recognition question (independently of the final answer participants arrive at) may sometimes foster false beliefs, we explored whether some participants would switch from an initially correct answer in the recognition task to accepting the misinformation in the SM task if they had the opportunity. As such a shift was not possible in the SM task we had used so far, a new version of the SM task was added in Experiment 3.

Experiment 3

Method

Participants. In Experiment 3, 162 participants were tested. Overall, 40 of them had to be excluded due to not remembering the warning (10), problems with the video presentation (5), disturbances during the video presentation or major interruptions during the experiment (12)[9], or not believing at all that the information was actually generated by a former participant (13). The final groups consisted of 26 (misinformed and unwarned, old SM task), 26 (misinformed and unwarned, new SM task), 20 (misinformed and warned, old SM task), 24 (misinformed and warned, new SM task), and 26 (control, old SM task) participants. Mean age was *M* = 22.27 years (*SD* = 3.95), 96 participants were women, and only 3 participants were not undergraduate psychology students. Data were collected at a university abroad.

Design. The design of Experiment 3 was identical to Experiment 1. One warned and one unwarned group worked though the new SM task, whereas the other three groups completed the old SM task. All groups were included in the statistical analyses of the recognition task; the results of the SM task were examined separately according to the old and new measurement.

Material and procedure. Beyond the SM version used in the former two experiments (*old SM task*), a new version (*new SM task*) was introduced in this experiment. In this new task, participants read statements such as *"The victim was kicked when he was lying on the ground"* (includes the misinformation on the central aspect),*"The 'passive' perpetrator wore a hood pulled over his head"* (includes the misinformation on the peripheral aspect), or *"The 'passive' perpetrator was clean-shaven"* (refers to a comparison item). For each statement, participants had to decide on a scale with the anchors *1 NO, certainly not*, *5 unsure*, and *9 YES, certainly* whether the statement corresponded (a) to what they had seen in the video, and (b) to what their partner had reported (see Chambers & Zaragoza, 2001, for a similar

[9] For instance, four participants had to be excluded due to a fire practice, another four because of an interruption while they were watching the video.

approach). Apart from the two experimental groups that completed the new instead of the old SM task, the procedure and the material were identical to that of Experiment 1.

Results

Data from 122 participants were included in statistical analyses. For an overview of error rates in the different groups, see Table B1. The adjusted alpha level for additional analyses was .025.

Baseline error. The baseline error in the recognition task did not differ significantly between the central item (15.4%) and the peripheral item (34.6%), difference = 19.2%, 95% CI [-3.6, 41.8], $z = 1.67$, $p = .096$, $h = 0.59$, binomial test $p = .180$. The same applied to the SM task with 11.5% (central item) and 15.4% (peripheral item), difference = 3.8%, 95% CI [-15.1, 23.7], $z = 0.45$, $p = .655$, $h = 0.20$, binomial test $p = 1.000$.

Recognition task. As in the former experiments, a significant association was observed between the groups and the error rate in the central item, $\chi^2(2, N = 122) = 14.40$, $p = .001$, $\phi_c = .34$. Individual comparisons revealed a significant misinformation effect (increase from 15.4% to 57.7%; difference = 42.3%, 95% CI [20.2, 59.9], $z = 4.30$, $p < .001$, $h = 0.92$, FET $p = .001$). However, no significant effect of a warning was found this time (decrease from 57.7% to 56.8%; difference = 0.9%, 95% CI [-20.1, 22.7], $z = 0.09$, $p = .931$, $h = 0.02$), and the misinformation effect was not eliminated in the warned group (difference = 41.4%, 95% CI [19.0, 59.9], $z = 4.03$, $p < .001$, $h = 0.90$, FET $p = .001$).

The association was also significant for the peripheral item, $\chi^2(2, N = 122) = 11.77$, $p = .003$, $\phi_c = .31$. The error rate was significantly higher in the misinformed and unwarned (73.1%) than in the control group (34.6%), difference = 38.5%, 95% CI [14.6, 58.3], $z = 3.44$, $p = .001$, $h = 0.79$. As in the other experiments, the warning did not significantly reduce the error rate in the peripheral item (decrease from 73.1% to 68.2%; difference = 4.9%, 95% CI

[-14.1, 24.6], z = 0.52, p = .600, h = 0.11), and the error rate was significantly higher in the misinformed and warned group (68.2%) than in the control group (34.6%), difference = 33.6%, 95% CI [9.2, 55.3], z = 2.87, p = .004, h = 0.68.

Overall, participants committed fewer errors with the central item (57.3%) than with the peripheral item (70.8%) in the misinformed groups at a descriptive level, difference = 13.5%, 95% CI [-0.8, 27.6], z = 1.86, p = .063, h = 0.27.

SM task. For these analyses, only the old SM task was considered. As with the recognition task, a significant association was found between the groups and the central item, $\chi^2(2, N = 72) = 12.82$, p = .002, ϕ_c = .42. The misinformation effect persisted (increase from 11.5% to 53.8%; difference = 42.3%, 95% CI [17.4, 63.0], z = 3.64, p < .001, h = 0.95, FET p = .003), and, as with the recognition task, the warning had no effect (increase from 53.8% to 55.0%; difference = 1.2%, 95% CI [-28.0, 30.8], z = -0.08, p = .938, h = -0.02, FET p = 1.000). The error rate was significantly higher in the misinformed and warned group (55.0%) than in the control group (11.5%), difference = 43.5%, 95% CI [15.0, 66.9], z = 3.40, p < .001, h = 0.98, FET p = .003. For the peripheral item, the chi-square test was not significant, $\chi^2(2, N = 72) = 4.89$, p = .087, ϕ_c = .26, FET p = 1.000.

In contrast to the recognition task, the error rate was significantly higher for the central item (54.3%) than the peripheral item (23.9%), difference = 30.4%, 95% CI [9.9, 49.7], z = -2.86, p = .004, h = -0.62, binomial test p = .007.

New source monitoring. The new SM task was introduced to check whether participants sometimes come to believe that they saw the suggested item in the original event although they had answered the recognition question correctly. The criterion for such a change was that participants select a value of 6 or more on the 9-point Likert scale, which means that they think they may remember the misinformation from the video. For the central item, none of the 26 participants who answered the recognition question correctly (misinformed warned and unwarned groups collapsed) changed his/her mind. With the peripheral item, only one (7.7%) out of the 13 participants altered his/her view and indicated that he/she might have seen the pulled-up hood in the video.

Discussion

The misinformation effect found in Experiment 2 was replicated in Experiment 3 for both items in the recognition task. In contrast to Experiments 1 and 2, however, the warning effect could not be corroborated for the central item. Thus, no evidence was found that the false reports of the central item were at least partially due to deliberation or the recency bias. We compared the sample from Experiment 3 with the other two samples, but we could not find any differences that would have explained this finding, in our opinion. It is possible that the effect did not occur because of a sampling error or because the warning has different effects in different populations.[10] As a logical consequence of finding a misinformation effect but no warning effect, the misinformation effect was also observed in the misinformed and warned group in both items. While the misinformation effect persisted in the SM task for the central item, it failed to reach significance for the peripheral item, even though the error rate was nearly twice as high in the misinformed and unwarned group as in the control group.

What could be replicated is the finding that the error rates for the two items differ depending on the memory task that is considered. The error rate was descriptively higher in the peripheral item when the recognition task was considered, but this was reversed in the

[10] The fact that the low-credibility condition (see footnote 1) was not realised in this study is most probably not an explanation, as the warning effect was not systematically stronger in the low-credibility than in the high-credibility condition.

SM task, with participants committing over 30% more errors with the central item than the peripheral item. Hence, the results corroborate the assumption that the reasons for false reports are quite different for the peripheral and central items despite the warning having no effect in this case.

Finally, analyses of the new SM task did not provide evidence that participants change their mind after a correct recognition decision. Thus, answering the recognition task did not seem to foster source misattributions in general, at least not after a short delay of a couple of minutes at most. Therefore, the surprising warning effect in the SM task found in Experiment 2 (and by trend also in Experiment 1) still demands further explanation. We will return to this in the Discussion in the main article.

Considering the three experiments, it becomes clear that the results were not always unanimous. This may at least partially be explained by the statistical methods and the rather low power associated with them. We therefore collapsed the data from the three experiments for the overall analysis, which is reported in the main article. This allows a more precise and robust estimate of the actual effects.

Were the differences between the experiments due to sampling error?

The data from the three experiments were more closely analysed in order to assess whether the differences between the experiments were due to sampling errors.

Method

To examine this question, we used an approach that compares multiple contingency tables as described by Bortz, Lienert, and Boehnke (2008, pp. 158–159). More precisely, we tested whether the combinations of the different categories occurred comparably often in the samples from the different experiments. Our data results in 6x3 tables. Six because there were three conditions per experiment and in each condition the answer to the critical question could be correct or incorrect (3x2 = 6), and three because the data from the three experiments were compared. Overall, there were 4 of these 6x3 tables: one for the central item with the recognition task, one for the central item with the SM task, one for the peripheral item with the recognition task, and one for the peripheral item with the SM task. Please note that the data from post-test 1 was added on the data from Experiment 1 as the comparison only works for studies with identical conditions.

Results

The Bonferroni correction resulted in a critical *p* value of .013. None of the outcomes of the homogeneity analysis reached significance (central item, recognition task: $\chi^2(10, N = 429) = 21.47$, $p = .018$, $\phi_c = .07$; central item, SM task: $\chi^2(10, N = 379) = 16.00$, $p = .100$, $\phi_c = .06$; peripheral item, recognition task: $\chi^2(10, N = 429) = 11.23$, $p = .340$, $\phi_c = .05$; peripheral item, SM task: $\chi^2(10, N = 379) = 10.42$, $p = .404$, $\phi_c = .05$).

Discussion

For the central item in the recognition task a tendency towards significant differences between the experiments was observed. The high chi-square value was mainly due to the higher number of errors in Experiment 3 compared to Experiments 1 and 2 for the misinformed and warned group (which is why no warning effect was observed in Experiment 3). It is worth discussing the failure to replicate the warning effect in Experiment 3, as this study was run at another university in another country (for further comments on this issue, see the main manuscript). With respect to the absence of a misinformation effect in Experiment 1, we conclude that this was probably due to a sampling error.

Additional References

Ackil, J., & Zaragoza, M. (1998). Memorial consequences of forced confabulation: Age differences in susceptibility to false memories. *Developmental Psychology, 34,* 1358–1372. doi:10.1037/0012-1649.34.6.1358

Bortz, J., Lienert, G. A., & Boehnke, K. (2008). *Verteilungsfreie Methoden in der Biostatistik* [Non-parametric methods in biostatistics] (2nd rev. ed.). Berlin, Germany: Springer.

French, L., Garry, M., & Mori, K. (2011). Relative–not absolute–judgments of credibility affect susceptibility to misinformation conveyed during discussion. *Acta Psychologica, 136,* 119–128. doi:10.1016/j.actpsy.2010.10.009

Roediger, H. L. III, & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 803–814. doi:10.1037/0278-7393.21.4.803

Roediger, H. L. III, Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review, 8*, 385–407. doi:10.3758/BF03196177