

Explicit warnings reduce but do not eliminate the continued influence of misinformation

ULLRICH K. H. ECKER, STEPHAN LEWANDOWSKY, AND DAVID T. W. TANG
University of Western Australia, Crawley, Western Australia, Australia

Information that initially is presumed to be correct, but that is later retracted or corrected, often continues to influence memory and reasoning. This occurs even if the retraction itself is well remembered. The present study investigated whether the continued influence of misinformation can be reduced by explicitly warning people at the outset that they may be misled. A specific warning—giving detailed information about the continued influence effect (CIE)—succeeded in reducing the continued reliance on outdated information but did not eliminate it. A more general warning—reminding people that facts are not always properly checked before information is disseminated—was even less effective. In an additional experiment, a specific warning was combined with the provision of a plausible alternative explanation for the retracted information. This combined manipulation further reduced the CIE but still failed to eliminate it altogether.

There is abundant evidence that human memory is reconstructive (see, e.g., Roediger & McDermott, 1995) and that those reconstructions can induce errors into our recollections. For example, people readily form “memories” for information that was never presented but merely hinted at. When presented with a list of words such as BED, REST, AWAKE, TIRED, DREAM, . . . , roughly half the participants will recall the nonpresented word SLEEP, due to its strong association with the studied items (Deese, 1959; Roediger & McDermott, 1995; Seamon et al., 2002). There is also much evidence that people’s memories, once formed, are susceptible to postevent suggestions. For example, witnesses of a traffic accident may remember a nonexistent traffic sign if one was mentioned after the event (e.g., if people were asked “did the red car slow down at the yield sign?” when, in fact, there was a stop sign (see Ayers & Reder, 1998; Loftus, 1975; Loftus & Palmer, 1974).

The present article deals with another source of potential memory errors—the failure to correct initially believed information that turns out to be false. Contrary to the ease with which false memories can be created and true memories altered, the *elimination* of memories for information that is later revealed to be false—we refer to this as *misinformation*¹—has proven to be considerably more difficult. Misinformation continues to affect behavior, even if people explicitly acknowledge that this information has been retracted, invalidated, or corrected (Ecker, Lewandowsky, & Apai, in press; Ecker, Lewandowsky, Swire, & Chang, 2010; Gilbert, Krull, & Malone, 1990; Gilbert, Tafarodi, & Malone, 1993; H. M. Johnson & Seifert, 1994, 1998; Seifert, 2002; van Oostendorp, 1996; van Oostendorp & Bonebakker, 1999; Wilkes & Leatherbarrow, 1988; Wilkes & Reynolds, 1999).

For example, H. M. Johnson and Seifert (1994) presented participants with a story about a fictitious warehouse fire, allegedly caused by volatile materials stored carelessly in a closet. Participants were later told that the closet had actually been empty. Although participants later remembered this retraction, they still used the outdated misinformation to make inferences; for example, people might argue that the fire was particularly intense because of the volatile materials or that an insurance claim may be refused due to negligence. H. M. Johnson and Seifert (1994) termed this reliance on misinformation the *continued influence effect* (CIE). The CIE is robust and occurs in a variety of contexts, regardless of the particular story being presented and regardless of the test applied (Ecker et al., in press; H. M. Johnson & Seifert 1994, 1998; Wilkes & Reynolds, 1999).

The continued influence of misinformation is also detectable in real-world settings. For example, during the 2003 invasion of Iraq, the public was exposed to countless hints that weapons of mass destruction (WMDs) had been discovered in Iraq. Even though no such report was ever confirmed, these constant hints were powerful enough to engender, in a substantial proportion of the U.S. public, a longstanding belief in the presence of WMDs that has persisted, even after the nonexistence of WMDs became fully evident (Kull, Ramsay, & Lewis, 2003; Lewandowsky, Stritzke, Oberauer, & Morales, 2005). Unconfirmed hints can thus engender false memories in the public (analogous to the “sleep” example presented at the outset) that resist subsequent correction (analogous to the warehouse fire example above).

People’s continued reliance on misinformation can have obvious adverse consequences in other settings relevant

to the public's welfare—for example, when jurors are instructed to disregard tainted evidence. In laboratory analogs of court proceedings, it has been repeatedly shown that mock jurors continue to rely on inadmissible evidence, even when they claim to have obeyed instructions to ignore it (Fein, McCloskey, & Tomlinson, 1997; Kassin & Sukel, 1997). In light of these adverse consequences, the question of how the CIE might be reduced or eliminated takes on particular urgency.

Previous Attempts to Eliminate Continued Influence

The CIE typically has been explained by reference to a mental-event model that people build when trying to understand an unfolding event (H. M. Johnson & Seifert, 1994; van Oostendorp, 1996; Wilkes & Leatherbarrow, 1988). On this view, a retraction of central information creates a gap in the model, and—because people are apparently more willing to accept inconsistencies than they are voids in their event model—they continue to rely on misinformation. That is, people prefer to retain *some* information in crucial model positions (e.g., what caused something to happen or who was involved), even if that information is known to be discredited (H. M. Johnson & Seifert, 1994; van Oostendorp & Bonebakker, 1999).

Previous efforts to reduce the CIE have been pursued along various lines, most of which have remained unsuccessful. Those efforts have included attempts to increase the impact of a retraction (1) by wording it in a particularly clear and direct way (H. M. Johnson & Seifert, 1994), (2) by repeating it (van Oostendorp & Bonebakker, 1999), or (3) by supplementing the retraction with explanations that clarified the contradiction (i.e., giving reasons why the misinformation was first assumed to be factual but then discarded; Bush, Johnson, & Seifert, 1994). In other attempts to reduce the CIE, (4) participants were admonished to read the materials carefully (van Oostendorp, 1996), or (5) the retraction was presented immediately after the misinformation (rather than some time later) to limit its impact on the online construction of the event model (H. M. Johnson & Seifert, 1994; Wilkes & Reynolds, 1999). None of those attempts were successful.

To date, only two factors have been identified that can reliably reduce reliance on misinformation: First, the effect is reduced—but rarely is eliminated—if the correction goes beyond a mere retraction by also providing a *causal alternative*. Thus, if a fire is initially linked to careless storage of volatile materials, a later correction stating that no such materials had been found will be particularly effective (i.e., reduce references to negligence in subsequent recall and inference tests), if the correction additionally states that arson-related materials were discovered (Ecker et al., in press; H. M. Johnson & Seifert, 1994; Rapp & Kendeou, 2007; but see van Oostendorp & Bonebakker, 1999). In terms of the mental-event model theory, the alternative can be used to fill what would otherwise be left void by the retraction in the event model.

Unfortunately, however, there are many real-life situations in which plausible alternatives may not be available (e.g., in a court case, there may not be an alterna-

tive suspect). Furthermore, H. M. Johnson and Seifert (1999) demonstrated that the causal alternative needs to be provided within the event context. They instructed participants to self-generate causal alternatives after reading the retraction, but this did not make the retraction more efficient. Thus, it appears that people do not revise their event model and do not replace the misinformation with alternative information, even when those alternatives are demonstrably accessible.

The second factor that seems to reduce the CIE is *suspicion* toward the source of the misinformation. In the WMD studies discussed earlier, belief in the existence of WMDs in Iraq was correlated with support for the war and was especially pronounced in those people who obtained news from sources that supported the invasion (e.g., Fox News; Kull et al., 2003). Lewandowsky et al. (2005) uncovered a more direct link between suspicion and the ability to update misinformation related to the Iraq War. They operationalized suspicion as the extent to which respondents doubted the official WMD-related reasons for the invasion. Lewandowsky et al. (2005) found that, when this measure was used as a predictor variable, it explained nearly a third of the variance in people's belief in misinformation. Moreover, once suspicion was entered as a predictor, previously striking mean differences between respondents in the U.S. and two other countries (Germany and Australia) disappeared and were, instead, found to reflect differing degrees of suspicion between those countries. Lewandowsky, Stritzke, Oberauer, and Morales (2009) extended the notion of suspicion by suggesting that it may be related to a more stable personality trait of skepticism—skeptics will generally tend to question the motives behind the dissemination of information.

Likewise, in mock jury studies, participants' reliance on inadmissible evidence was reduced if, immediately after presentation of the inadmissible evidence (e.g., libelous newspaper reports or hearsay testimony), participants were reminded that the media may publish biased reports to raise sales or that hearsay evidence may be introduced to deliberately manipulate the jury's thinking (Fein et al., 1997).

The literature thus suggests that suspicion may be capable of reducing the CIE. However, suspicion will be useful in reducing the CIE only in situations in which people believe that there are reasons to be suspicious in the first place, and, in many situations, it will not be feasible to plausibly induce suspicion. Moreover, as we discussed earlier, the effectiveness of induced suspicion may be moderated by a person's level of skepticism, which may represent a stable personality trait (Lewandowsky et al., 2009); hence, it is difficult to manipulate. Given these constraints, we investigated a novel, but related, avenue to reduce people's reliance on misinformation by examining the effects of explicit prior warnings about potentially misleading information. Participants in the present study received one of two types of warning. One was designed to induce a general sense of *alertness* that, frequently, information first presented as factual is later retracted. We believe that this type of induced alertness is similar to induced suspicion (cf. Fein et al., 1997), but without requir-

ing people to assume that someone is willfully trying to mislead them. The other type of warning was designed to additionally provide specific information about the CIE, on the assumption that explaining the exact nature of the effect should enable participants to avoid it.

The notion that direct warnings featuring specific information can counteract memory and reasoning errors has found some support in previous research. Like CIEs, false memories in the Deese/Roediger–McDermott paradigm have proven to be extremely difficult to eliminate. For example, even after five study–test trials (of the same list), the rate of false recall of unstudied words (e.g., SLEEP) after studying strong associates (e.g., DREAM, BED, . . .) has been found to be as high as 32% (McDermott, 1996). This effect can be reliably reduced, however, by telling participants about the effect before list presentation (Gallo, Roberts, & Seamon, 1997; Jou & Foreman, 2007; McCabe & Smith, 2002; McDermott & Roediger, 1998; Peters et al., 2008; Watson, McDermott, & Balota, 2004). Warnings presented after list presentation and prior to recall are less effective (Anastasi, Rhodes, & Burns, 2000; Gallo, Roediger, & McDermott, 2001; McCabe & Smith, 2002; Neuschatz, Payne, Lampinen, & Tolia, 2001).

Similarly, the power of postevent suggestions in the Loftus paradigm (e.g., suggesting that there was a yield sign when there was not) can be abated by warning participants before the suggestion is administered (Chambers & Zaragoza, 2001; Greene, Flynn, & Loftus, 1982). The effect of postsuggestion warnings appears more mixed, with some authors reporting a reduction of false memories (Chambers & Zaragoza, 2001; Christiaansen & Ochalek, 1983; Echterhoff, Hirst, & Hussy, 2005; Zaragoza & Koshmider, 1989; see also Lindsay & Johnson, 1989) and others not (Belli, Lindsay, Gales, & McCarthy, 1994; Greene et al., 1982). This state of affairs suggests that warnings alter processing, both at encoding (e.g., by leading participants to infer what the suggested concept is in a false-memory paradigm) and at retrieval (e.g., by boosting source-monitoring efforts).

Should explicit warnings prove to be able to reduce the CIE, it would be interesting for both applied and theoretical reasons. In applied settings, briefings of jurors and media education measures could be used to reduce the unwanted effects of outdated information on people's inferences. From a theoretical perspective, the effects of warnings could shed more light on the mechanisms underlying misinformation effects. For example, it has been suggested that misinformation effects result mainly from automatic retrieval processes—that is, inadequate reliance on active, but invalid, information paired with insufficient retrieval control (Ayers & Reder, 1998; H. M. Johnson & Seifert, 1994; M. K. Johnson, Hashtroudi, & Lindsay, 1993; Schwarz, Sanna, Skurnik, & Yoon, 2007). Warnings might lead to an increase in strategic retrieval monitoring and should therefore prove effective in reducing the CIE.

We present two experiments that explored the effects of general and more specific warnings on the CIE. Given previous research, we chose to present warnings before the encoding of the relevant materials to maximize their effectiveness. Materials consisted of a fictitious accident

report that contained a correction of an initial false statement and a subsequently administered inference and fact-recall questionnaire.

EXPERIMENT 1

Experiment 1 used a standard continued influence paradigm (cf. H. M. Johnson & Seifert, 1994), featuring four retraction conditions plus a baseline no-retraction control condition in a between-subjects design. Participants read a fictitious account of a minibus accident whose victims were initially said to be elderly people. In the four retraction conditions, this information was later revoked. The retraction conditions differed in terms of whether the story contained a warning at the outset or contained an alternative account of who the passengers were later in the story. Participants were ultimately given a questionnaire to probe their understanding of the story and the extent of their reliance on the misinformation.

Method

Participants. One hundred twenty-five undergraduate psychology students (mainly first-year students; 30 males; mean age, 19.1 years) participated for partial course credit. Participants were tested individually and were assigned randomly to the five conditions ($n = 25$ in each condition).

Stimuli. Participants were presented with a folder containing a series of 14 statements, which provided a fictitious account of a minibus accident. Each statement was printed on a separate sheet of paper. (Materials are provided in Appendix A.)

In all conditions, Message 4 provided the (mis)information that the passengers were elderly: “Police have stated that the passengers on the bus were a group of elderly people who were on their way back to their nursing home.” In the no-retraction control condition, this initial information was not challenged, thus providing a baseline measure for the final inference test. In all four retraction conditions, the misinformation was retracted in Message 11: “Police stated that the passengers on the minibus were not elderly people.” In the retraction-only condition, this was the only correction. In the alternative condition, the retraction was accompanied by a plausible alternative piece of information within the same message: “Passengers on the minibus were not elderly people but college hockey players returning from a victory party.” This condition aimed to replicate the effect of providing an alternative for the retracted misinformation (H. M. Johnson & Seifert, 1994) and to compare the effects of warnings with the gold standard of previous research.

In the two remaining retraction conditions, only the simple retraction was presented in Message 11, but participants additionally received a written warning before reading the scenario. In the general warning condition, the warning stated that sometimes reported “facts” are not double-checked before they are released (see Appendix B for the exact wording of both warnings). Finally, in the specific warning condition, the warning explained the CIE very specifically and provided two concrete examples of its operation.

Procedure. Participants read the statements at their own pace without backtracking, although encoding time was limited to 5 min (none of the participants approached this limit). Participants were aware of an upcoming memory test. They then engaged in an unrelated 15-min distractor task before responding to a questionnaire comprising 9 inference questions, 9 fact-recall questions, and 2 manipulation-check questions. All questions are included in Appendix C.

Inference questions were designed to elicit inferences about aspects of the incident that were directly or indirectly linked to the misinformation—viz. the age of the passengers. For example, the question *Why do you think it was difficult getting both the injured and*

uninjured passengers out of the minibus? could be answered by referring to the passengers' old age, although one could also come up with alternative explanations, such as—in the alternative condition—the hockey players being intoxicated after their victory party.

The fact-recall questions were designed to check for adequate encoding and to assess overall memory for details of the scenario, but were not directly related to the age of the passengers. The final recall question was an exception, because it directly asked who was on the minibus.

The manipulation-check questions targeted participants' awareness and understanding of the retraction. The questionnaire items were always given in this order, to avoid carryover effects from recall and from the manipulation check to the inference part.

Results

Coding procedure. Following a standardized scoring guide, a trained scorer, who was blind to experimental conditions, scored the questionnaires. Another scorer scored 20 randomly selected questionnaires (4 from each condition) for an interrater reliability check. Reliability was found to be high ($r = .97, .84, \text{ and } .94$ for inference, fact, and manipulation-check questions, respectively).

Analysis focused on three dependent measures: participants' reliance on misinformation (inference score), the accuracy of their recall (fact-recall score), and their acknowledgment of the retraction (manipulation-check score). For the inference score, any reference to elderly people or items clearly related to the elderly stereotype incremented the inference score. Examples of such inferences include "*The people found it difficult to exit the bus because they were frail and weak*" or "*The passengers' children were informed about the accident.*" Inferences that were contained in fact-recall or manipulation-check responses also contributed to the inference scores; hence, the maximum inference score was 20 (1 possible inference per item).

Fact questions were given a recall score of 1 for a correct response and 0 for an incorrect response. The fact-recall responses were used to ensure that all participants had read and understood the story; no one scored below 2 (out of 8),² and, thus, all participants were retained for analyses.

Manipulation-check questions were given a score of 1 if participants remembered a retraction and 0 if they did not. The maximum score was 2.

Inferences. The mean numbers of references to misinformation in all conditions are shown in Figure 1. As expected, the no-retraction control condition had the highest score; the specific warning and alternative conditions had the lowest scores.

Instead of an undirected omnibus ANOVA that would have required multiple follow-up tests, we used our a priori hypotheses to guide a contrast analysis, following Rosenthal, Rosnow, and Rubin (2000). The results are summarized in Table 1.

First, we specified a set of contrasts that reflected our coarse initial hypothesis, based on previous research, that a mere retraction would not significantly reduce the number of references to misinformation, whereas a general warning would moderately reduce continued influence, and a specific warning or provision of an alternative

would maximize the reduction. We, therefore, assigned lambda weights 1, 1, 0, -1 , and -1 to the no-retraction, retraction-only, general warning, specific warning, and alternative conditions, respectively. This contrast (C1 in Table 1) yielded a highly significant (directed) main effect of condition.

Further planned contrasts showed that a mere retraction reduced the number of references to misinformation numerically, but this was not significant at the conventional .05 alpha level (C2 in Table 1). This is in line with previous research, which has variously found significant effects of retractions (e.g., Wilkes & Reynolds, 1999) as well as clear null effects (e.g., H. M. Johnson & Seifert, 1994); the reasons for this inconsistency remain unclear and are outside of the scope of this article. Second, an additional general warning was not capable of reducing the level of continued influence found with a mere retraction (C3). Finally, both a specific warning and the provision of an alternative were not only able to strongly reduce the CIE, but, in this regard, also proved to be more efficient than the general warning was (C4, C5).

It is noteworthy that the reduction of the CIE by a specific warning or alternative occurred quite uniformly across all items carrying the effect. References to misinformation were made mainly in response to Inference Questions 1–5, and Fact Question 9 (these items had between 14 and 24 references in the no-retraction condition, summed over all participants; all other items scored below 10). The alternative and the specific warning reduced references to misinformation in response to these items by 51%–65% (42%–67% in the alternative condition, 53%–63% in the specific warning condition).

Further analyses dealt with the alternative interpretation that our warnings did not specifically reduce references to misinformation but rather overall output more generally (cf. Goldsmith, Koriat, & Weinberg-Eliezer, 2002). To address this possibility, a naive scorer quantified participants' responses to the inference questionnaires by means of a propositional analysis (i.e., a count of "idea units"; Kintsch & Greene, 1978) and also a more fine-grained count of *lexical information units* (LIUs; i.e., all informative, well-formed content and function words that carry semantic meaning and are relevant to the narrative; Marini, Boewe, Caltagirone, & Carlomagno, 2005). We then compared the output of the retraction-only group with those of the two warning groups (because all three conditions featured a retraction and differed only in the administration of a warning). The data are summarized in Table 2. To test the directed hypothesis that warnings reduced overall output, we contrasted the retraction-only condition with the combined warning conditions (lambda weights 2, -1 , -1), but found no significant effects on the mean number of propositions or LIUs (both $F_s < 1$). Hence, we found no reduction in output quantity in the warning groups relative to the retraction-only group. It follows that warnings reduced the CIE without reducing either general output or—as will be shown next—correct recall.

Recall. We did not expect our manipulations to affect fact recall. Mean fact-recall scores across all conditions ranged from 4.92 to 5.40 (out of 8). A one-way ANOVA

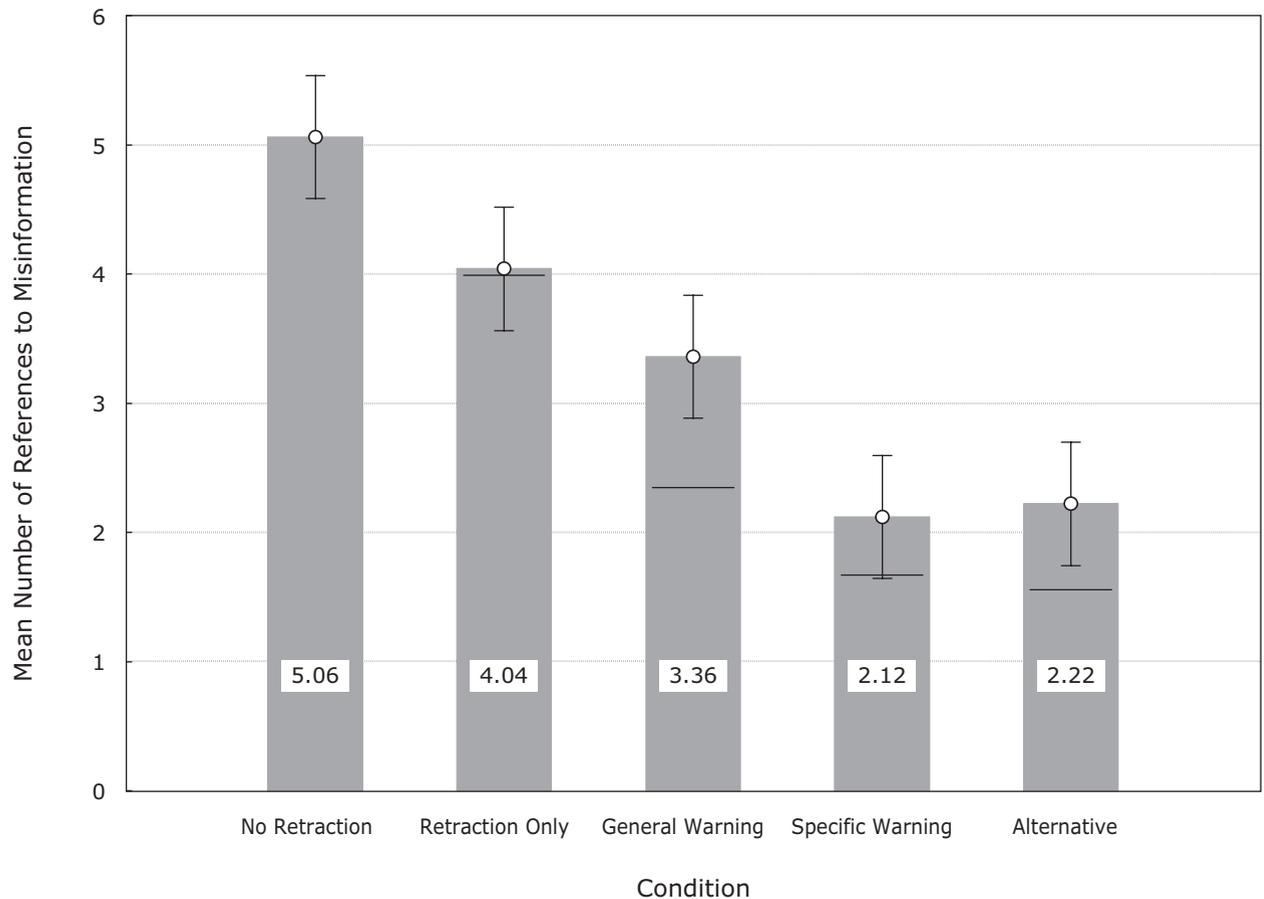


Figure 1. Mean number of references to misinformation across experimental conditions in Experiment 1. Each data point is based on 25 participants. The maximum score is 20. Error bars show standard errors of the means. Horizontal lines inside the bars of the four conditions featuring a retraction indicate mean reference-to-misinformation scores based on subsamples of participants remembering the retraction; these data points are based on 19, 17, 20, and 18 participants (left to right).

revealed no main effect of condition ($F < 1$), thus obviating the need for a contrast-based analysis.

Awareness of retraction. Mean manipulation-check scores across conditions (excluding the no-retraction control condition, for which the questions were meaningless) ranged from 1.00 to 1.24 (out of 2). A one-way ANOVA failed to reveal a main effect of condition ($F < 1$).

Next, we reanalyzed the inference results after removing 32 participants who scored 0 on the manipulation-check questions and, hence, may not have noticed the retraction in the first place. Inference scores of this restricted sample are also shown in Figure 1, and results of the contrast analysis are shown in Table 1. The overall pattern was similar, and there was still a main effect of condition in the directed contrast analysis (C1 in Table 1). However, in this subsample, the general warning was apparently more effective in reducing the CIE; that is, there was now a significant difference in the number of references to misinformation between the general warning and the retraction-only conditions (C3), but there was no longer a difference between the general warning versus the specific warning and alternative conditions (C5).

Discussion

Experiment 1 replicated previous research in several important ways. First, we again showed that a mere retraction can be insufficient to significantly reduce continued reliance on misinformation (H. M. Johnson & Seifert, 1994; but see also Bush et al., 1994; Ecker et al., 2010; Wilkes & Reynolds, 1999). The CIE persisted, even when only those participants were considered who remembered the correction, implying that the continued influence of misinformation was not due to a trivial failure to encode or to retain the retraction. Second, if the retraction was accompanied by a plausible alternative, participants were able to reduce their reliance on misinformation (Ecker et al., 2010; H. M. Johnson & Seifert, 1994; Rapp & Kendeou, 2007). The fact that we replicated those two pervasive aspects of previous research permits us to interpret the novel results with greater confidence.

The novel finding of Experiment 1 was that a specific warning—by providing a detailed account of the CIE—reduced people's reliance on misinformation. In doing so, the specific warning was as successful as the provision of an alternative account—the principal factor hitherto

Table 1
Contrasts Calculated on Inference Scores in Experiment 1

Contrast	Lambda Coefficients for Conditions					Full Sample		Restricted Sample	
	No Retraction	Retraction Only	General Warning	Specific Warning	Alternative	<i>F</i> (1,120)	<i>p</i>	<i>F</i> (1,88)	<i>p</i>
	C1	1	1	0	-1	-1	24.91	<.0001	29.51
C2	1	-1	0	0	0	2.29	.13	1.48	.23
C3	0	1	-1	0	0	1.02	.32	5.23	.025
C4	0	2	0	-1	-1	10.25	.002	13.95	<.001
C5	0	0	2	-1	-1	4.15	.04	0.97	.33

Note—Restricted sample consists of participants who did recall the retraction.

known to effectively reduce the CIE. A general warning designed to induce nonspecific alertness was effective only in participants who also explicitly remembered the retraction. This implies that only participants who became fully aware of the retraction during encoding were able to associate it with the initial warning. We discuss this pattern more extensively in the General Discussion section.

We also showed that, even with a specific warning and even when an alternative was provided, the CIE persisted. None of the manipulations used in Experiment 1 were sufficient to eliminate the continued influence of misinformation entirely. This is also in line with previous research: H. M. Johnson and Seifert (1994) reported that the provision of an alternative eliminated the CIE in their “thematic inference” measure (comprising only references to misinformation made on their inference questionnaire) but not in their “direct reference” measure, which additionally included references to misinformation made in their tests of recall (and which is hence comparable to the misinformation measure used here). Ecker et al. (in press) reported three experiments featuring a causal-alternative condition and found persistent residual reliance on misinformation in all three, as did van Oostendorp and Bonebakker (1999).

In the next experiment, we asked whether, by combining the two most powerful manipulations known—provision of an alternative plus a specific warning—we can entirely eliminate reliance on misinformation.

EXPERIMENT 2

According to the mental-event model account of the CIE, an alternative serves to fill the gap in the event model left behind by a retraction. This explains both the failure of a direct, specific warning to completely eliminate continued influence and the efficiency of an alternative in reducing the effect. However, the question remains why significant levels of continued influence persist, even when an alternative is made available. Ayers and Reder (1998) and H. M. Johnson and Seifert (1998) have argued that representations of valid and invalid information coexist in memory and compete for activation. Hence, the retracted information is most likely not completely removed from the event model but, instead, either is tagged as invalid or is suppressed. In some instances, the remaining activation

of invalid information then seems sufficient to produce residual reliance on misinformation.

Experiment 2 was designed to examine whether alerting people to the effects of misinformation could eliminate the CIE in cases in which there is alternative information available (e.g., by boosting suppression of invalid information or by enhancing strategic retrieval monitoring; the likely mechanisms will be discussed more thoroughly in the General Discussion section). The experiment featured the same scenario as that in Experiment 1, using a between-subjects design with three conditions, all of which featured the specific warning at the outset. The main condition of interest combined the specific warning with the provision of an alternative—that is, the most effective manipulations of Experiment 1 (specific warning + alternative condition). The two remaining conditions simply served to provide upper and lower comparison baselines: In one condition, the passengers were elderly people throughout (the no-retraction condition, identical to that in Experiment 1, except that participants received the specific warning), whereas, in the other condition, the passengers were young hockey players throughout (the alternative-throughout condition).

Method

Participants. Ninety-two undergraduate psychology students (mainly first-year students; 26 males; mean age, 19.9 years) participated for partial course credit. None had taken part in Experiment 1. Participants were tested individually; the first 75 were assigned randomly to the three experimental conditions. We then conducted an initial analysis and decided to stock up the specific warning + alternative condition to ensure that the results were not driven by a small number of outliers and to administer an additional debriefing questionnaire. In this extension, a further 17 participants were tested in the specific warning + alternative condition (hence, there were 42

Table 2
Lexical Output in Experiment 1

Condition	Propositions		Lexical Information Units	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Retraction only	24.48	1.49	58.68	5.50
General warning	27.80	1.35	64.88	4.69
Specific warning	22.80	1.52	57.20	5.13

Note—*n* = 25 in each condition.

participants in this condition and 25 in each of the two control conditions). Results from the initial subsample ($n = 25$) are not reported separately but were in accordance with the results obtained with the full sample and are reported below.

Materials and Procedure. The materials and procedures used in this experiment were identical to those in Experiment 1, with two exceptions: The specific warning was presented in all conditions, and, in the alternative-throughout condition, the passengers were initially (Message 4) reported to be young instead of old, and no further correction was presented. For the extension sample in the specific warning + alternative condition, a four-question debriefing questionnaire was administered at the end of the experiment to more thoroughly assess understanding of and belief in the correction (Appendix C lists the questions).

Results

Coding procedure. The data were scored as they were in Experiment 1. Using the same inclusion criterion as before (fact-recall rate $> .25$), all participants were retained for analysis.

Inferences. The alternative-throughout condition was excluded from analysis, because it—quite expectedly—did not contain any references to the elderly theme, be-

cause this information was never presented (i.e., an inference score of zero). This implies that any references to misinformation made in the specific warning + alternative condition can be safely interpreted as a true CIE; they would not have been made without the misinformation being presented in the first place.³

Data from the two remaining conditions are shown in Figure 2. Scores in the specific warning + alternative condition were significantly lower than those in the no-retraction condition [$F(1,65) = 56.81, MS_e = 3.58, p < .0001$].

Of greatest interest is whether the remaining small extent of reliance on misinformation in the specific warning + alternative condition was significantly greater than zero. A one-sample t test confirmed this for the sample overall [$t(41) = 3.78, p < .001$] and also with the restricted sample of participants who recalled the retraction [i.e., with a manipulation-check score $> 0; n = 37, M = .84, SE = .26; t(36) = 3.17, p < .01$].⁴

Altogether, almost half of the participants in the specific warning + alternative condition made at least one

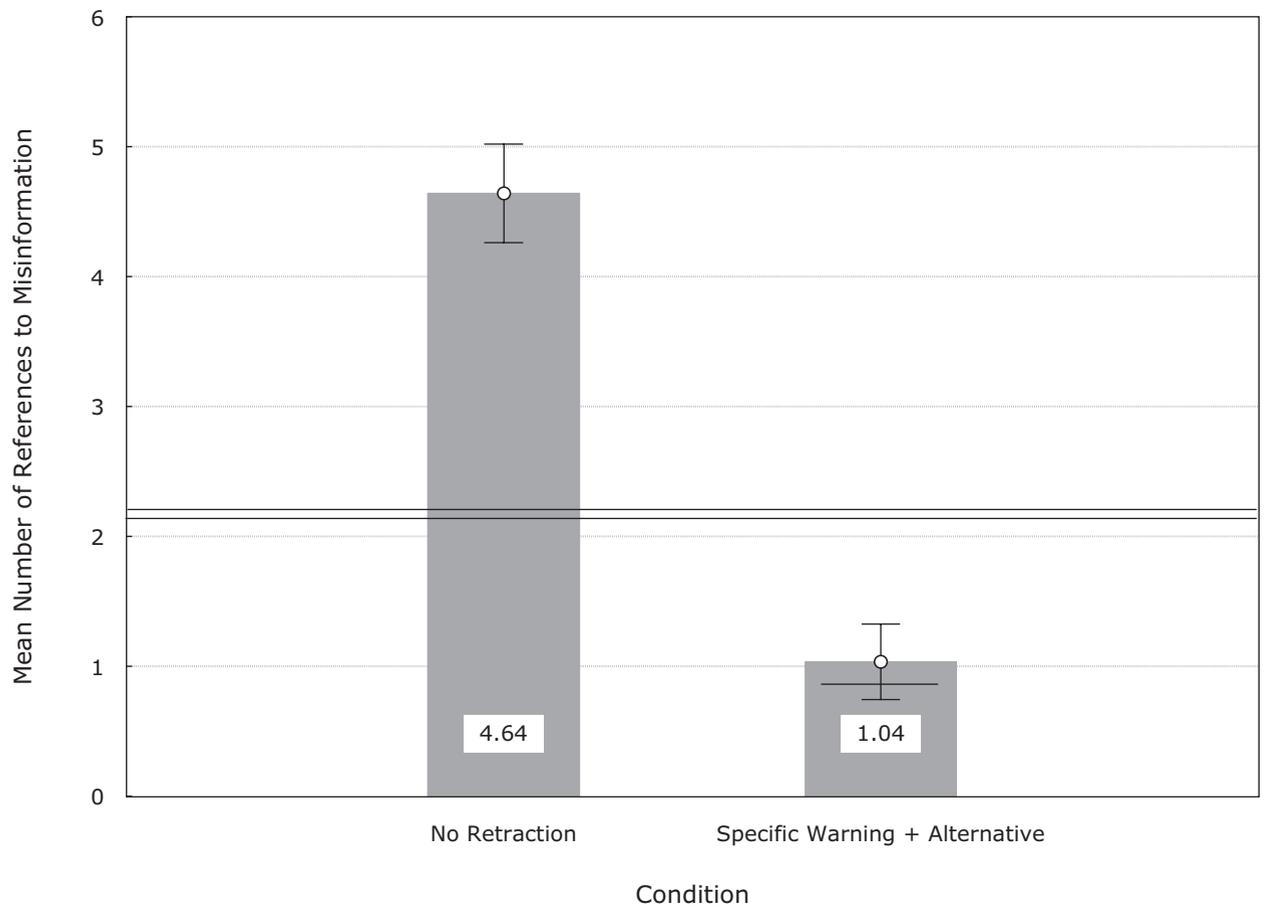


Figure 2. Mean number of references to misinformation across experimental conditions in Experiment 2 (note that the mean of the alternative-throughout control condition was zero). The data points of the no-retraction and specific warning + alternative conditions are based on 25 and 42 participants, respectively. Error bars show standard errors of the means. The horizontal line inside the specific warning + alternative bar indicates the mean reference-to-misinformation score based on the subsample of participants remembering the retraction; this data point is based on 37 participants. To allow across-experiment comparison, the two horizontal lines spanning the entire graph refer to the means of the specific warning (lower line) and alternative (upper line) conditions of Experiment 1.

reference to misinformation (17 out of 42). The debriefing questionnaire administered to the extension sample revealed that virtually all participants correctly considered the experiment to be about event memory and memory updating. Two participants demonstrated clear source confusion, in that they classified misinformation as correct information and vice versa. Some participants correctly identified the misinformation and the correction but indicated that they did not believe the correction. Interestingly, however, **the number of references to misinformation seemed completely independent of participants' belief in the correction**, with those participants expressing doubts about the validity of the correction ("I didn't believe it" and "not sure" responses) making as many references to misinformation ($n = 7$; $M = 1.29$; $SE = 0.53$) as did participants indicating that they did believe the correction ($n = 10$; $M = 1.30$; $SE = 0.72$). None of the participants indicated that they tried to "outsmart" the experimenter by purposely responding in accordance to misinformation because they thought the correction was a trick.

Comparing inference scores with those in Experiment 1. Finally, we compared the specific warning + alternative condition of Experiment 2 with the two corresponding individual conditions of Experiment 1 (specific warning and alternative, respectively). A planned contrast confirmed that the specific warning + alternative condition (of Experiment 2) differed statistically from the two conditions of Experiment 1 [$F(1,89) = 6.52$, $MS_e = 4.50$, $p = .01$].

Discussion

Experiment 2 showed that the combined effect of the specific warning and the provision of an alternative account reduced reliance on misinformation more than the constituent strategies alone. Notwithstanding, the CIE still persisted despite the combined power of both variables.

GENERAL DISCUSSION

In the present study, we investigated whether prior warnings can reduce the continued influence of misinformation. In replication of previous research, we found that a retraction alone failed to substantially reduce reliance on misinformation, even though participants acknowledged and remembered the retraction itself (H. M. Johnson & Seifert, 1994). Also in line with previous research, providing an alternative for the retracted misinformation strongly reduced but did not fully eliminate reliance on misinformation (cf. Ecker et al., 2010; H. M. Johnson & Seifert, 1994; Rapp & Kendeou, 2007).

The novel finding of our study was that a specific warning—that is, informing participants about the CIE up front in some detail—reduced reliance on misinformation after its retraction to about the same extent as did the provision of an alternative account. We also found that, by combining a specific warning with an alternative account, reliance on misinformation could be further reduced. Nonetheless, even when both of these manipulations were combined, misinformation lingered and a significant (although small) number of inferences were still present,

with more than a third of participants making at least one reference to corrected misinformation. A general warning also reduced the CIE, although this reduction was significant only in participants who demonstrably remembered the retraction. This result suggests that inducing alertness might be another effective way of reducing reliance on misinformation but that its effectiveness may be limited.

In contrast to the unsuccessful efforts to enhance the effectiveness of retractions (discussed in the introduction), the present study demonstrated that a large reduction in the CIE can be achieved by a simple but specific warning up front. The question then arises—how and why does a warning work?

Mechanisms That Mediate the Effects of Warnings

To our knowledge, this is the first study to examine the effects of warnings in the continued influence paradigm. By contrast, as reviewed at the outset, warnings have been used extensively in the Deese/Roediger–McDermott paradigm and also in the Loftus postevent suggestion paradigm. Various processes have been put forward by which warnings exert their protective effect, and these could prove informative for the present study.⁵

First, warnings could reduce potential demand characteristics (Eakin, Schreiber, & Sergent-Marshall, 2003; McCloskey & Zaragoza, 1985). Participants might have some naive theory about what is expected of them—for instance, that they should not believe unsubstantiated corrections—and this could lead to continued reliance on misinformation. Warnings could be effective because they clarify that participants are expected not to rely on outdated information. However, demand characteristics seem an unlikely source of the CIE. In a study by Rantzen and Markham (1992), misinformation was introduced either by a fellow participant or by the experimenter. The authors reasoned that social-demand effects should be higher in the latter case, but found no effect of this manipulation. Moreover, our debriefing questionnaire in Experiment 2 yielded no evidence that participants referred to misinformation in order to either please or outsmart the experimenter.

Second, the salience of misinformation at the time of retrieval could lead to output inhibition; that is, retrieval of misinformation might make retrieval of alternatives less likely (Eakin et al., 2003). In the case of retracted misinformation with no alternative provided, one could argue that salient misinformation could inhibit the spontaneous generation of alternative accounts at retrieval, leading to higher reliance on misinformation. Warnings may serve to counteract this retrieval inhibition, for instance, by enhancing alternative-generation efforts. Again, this appears rather unlikely: H. M. Johnson and Seifert (1999) found that participants who had received a retraction were able to generate as many alternative accounts as were control participants who were never presented with misinformation and that this generation of alternatives did not reduce the CIE in the first place.

Third, warnings could induce a shift in response criterion. If references to misinformation are made on the

basis of weaker evidence than are references to correct information, participants could become more accurate by becoming more conservative. However, according to the false-memory and postevent suggestion literature, warnings tend to affect only false alarms and not correct responses (Echterhoff et al., 2005; Gallo et al., 2001; Jou & Foreman, 2007; Peters et al., 2008; Watson et al., 2004). Also, if these false-memory phenomena were based on criterion setting at retrieval, it would be unclear as to why postencoding warnings are generally found to be ineffective (Anastasi et al., 2000; Gallo et al., 2001; Greene et al., 1982; Neuschatz et al., 2001; cf. also Peters et al., 2008). Consistent with those precedents, we likewise found no group differences in overall response output or fact recall: Warned participants produced as many ideas and recalled as many event details as did unwarned participants. This is inconsistent with a criterion-shift explanation.

Fourth, warnings could lead to selective rehearsal of correct information (Greene et al., 1982). In the present case, after a retraction, participants could focus on rehearsing the retraction. This seems rather unlikely, however, given Greene et al.'s findings from the Loftus paradigm that increasing the time between a warning (given immediately after the encoding of an event) and misleading suggestions did not reduce the subsequent effects of misinformation on a recognition test, and that warned participants did not remember valid information any better than unwarned participants did, even if there was no misinformation. Also, Eakin et al. (2003) found immediate postmisinformation warnings that explicitly identified the piece of misinformation (thus, allowing selective rehearsal of the corresponding correct piece of information) as effective as warnings that did not specify the misleading information.

Fifth, warnings could lead to more efficient encoding (e.g., Gallo et al., 2001; Neuschatz, Benoit, & Payne, 2003). Consistent with this notion, warnings given before encoding are typically more effective than are postencoding warnings. In the false-memory paradigm, it is likely that this encoding effect occurs selectively for the unrepresented "misinformation," because only the false alarm rate is lowered, with no effect on retrieval of actual study items (Gallo et al., 2001; Jou & Foreman, 2007; Peters et al., 2008; Watson et al., 2004). This pattern of findings is explained by an activation-monitoring account (cf. Neuschatz et al., 2003; Peters et al., 2008): Study list items activate the common associate, but a monitoring process can be used at encoding to "figure out" the nonpresented associate and then to consciously "tag" it as unstudied. Fittingly, Neuschatz et al. (2003) reported that warnings were effective mainly in preventing false memories for unstudied words if the concept was easy to identify given the list of studied associates. In contrast, a generally enhanced encoding (e.g., more strategic encoding of item-specific information; McCabe & Smith, 2002) should affect both hit and false alarm rates. To our knowledge, there is only one study reporting such an effect (Westerberg & Marsolek, 2006), but this study used many lists, presumably making the proposed tagging operation less efficient. In the continued influence paradigm, misinformation is

arguably also active when it is corrected, and, at that point in time, participants can presumably tag it as misinformation, which will facilitate its rejection at retrieval.

Sixth, Eakin et al. (2003) have suggested that warnings may lead to active suppression of the misinformation's activation. As mentioned before, representations of valid and invalid information may coexist in memory and compete for activation (Ayers & Reder, 1998; H. M. Johnson & Seifert, 1998). It follows that inhibition of misinformation should lead to prevalence of correct information. Eakin et al. supported their claims by demonstrating that warnings were less successful if misinformation was repeated; they argued that suppression could no longer be achieved if misinformation is strengthened via repetition. In the context of the CIE, suppression of misinformation could make it less salient and available and could, hence, indirectly facilitate retrieval monitoring. In support, findings that the effectiveness of postmisinformation warnings correlates positively with working memory capacity and negatively with age (McCabe & Smith, 2002; Watson et al., 2004) add weight to the assumption that active suppression processes (which should covary with working memory abilities) are invoked by warnings. However, these working memory correlations are also in line with the following, final alternative—namely, that warnings foster processing-intensive monitoring processes at retrieval.

This final alternative, which we consider to be the most convincing, assumes that warnings affect retrieval processes directly (McCabe & Smith, 2002; Watson et al., 2004). According to the source-monitoring account (M. K. Johnson et al., 1993; cf. Chambers & Zaragoza, 2001; Echterhoff et al., 2005; McCabe & Smith, 2002; Zaragoza & Lane, 1994), references to incorrect information stem from source confusion. For example, people could falsely remember that information given in a post-event suggestion was actually contained in the original event. An alternative, but complementary, account was proposed by Ayers and Reder (1998), who suggested that, in memory, chunks of information (including both "correct" and "incorrect" concepts) compete for activation; people are spontaneously aware only of activated concepts, whereas the links between concepts and their context (e.g., their source) need to be retrieved strategically. A warning might, therefore, lead participants to rely less on automatic processing at retrieval, thus guarding against the unfiltered use of strongly activated, but misleading, information (Schwarz et al., 2007; Whittlesea & Williams, 2001), and instead lead participants to engage in more strategic retrieval processing, with a focus on source monitoring. Consistent with this notion, in a postevent suggestion study, Echterhoff et al. (2005) reported that warned participants rated veridical event memories to be more vivid than did participants who received no warning, and they also took longer to reject incorrect information. Overall, the behavior of warned participants mimicked that of participants receiving explicit source-monitoring instructions. In further support of the source-monitoring notion, misinformation effects are usually reduced when source memory tests (rather than recall or recognition tasks) are

used, even without warnings (Lindsay & Johnson, 1989; Zaragoza & Lane, 1994; but see Belli et al., 1994, and, for a review, see Chambers & Zaragoza, 2001).

After a retraction of misinformation, rather than a misattribution of information to an incorrect source (e.g., witnessed vs. suggested), it is the outdated association of a piece of information with a causal role within a mental-event model that produces references to misinformation. Hence, it is this association to a role (rather than a specific source) that participants need to monitor at retrieval.

In summary, a warning most likely reduces the CIE by allowing participants to tag misinformation as outdated at encoding, putting participants in an altered retrieval mode (cf. Ayers & Reder, 1998). The altered retrieval mode boosts monitoring processes at retrieval (which, in turn, profit from tagging at study and active suppression of misinformation's activation; cf. Jou & Foreman, 2007; Peters et al., 2008).

Another interesting question is "Why was our specific warning more successful than our general warning?" By giving concrete examples, the specific warning made the temporal dynamics of the situation more explicit (misinformation . . . retraction . . . reliance on misinformation). Hence, one could argue that the specific warning may have been more successful during both encoding and retrieval. However, the fact that the general warning worked as well as the specific warning did in participants who acknowledged the retraction suggests that it is the encoding stage, in particular, in which the specific warning was more efficient. This would imply that the tagging of misinformation would profit most from specific instructions, whereas more nonspecifically induced alertness would mainly lead to more strategic monitoring at retrieval. Although this remains speculative, a connection between induced alertness and source monitoring is consistent with previous research using similar suspicion-inducing warnings (cf. Echterhoff et al., 2005, who used suspicion-inducing postencoding warnings and found that the performance of warned participants in a postevent misinformation paradigm mimicked that of participants explicitly instructed to do source monitoring).

Continued Influence of Misinformation—Unavoidable?

Wilson and Brekke (1994) argued that mental contamination—the unintentional impact of inappropriate information—occurs mainly because people are unaware of it and that warnings can help, inasmuch as they make people aware of a bias or contamination. This is obviously only the first step in a cascade, because people must also be able to counteract processing of the misinformation. This ability cannot be taken for granted: For example, Schul and Manzury (1990) showed that, when asked to judge a defendant's guilt, mock jurors who successfully avoided reliance on inadmissible information could not do so when asked to judge the defendant's aggressiveness or to give their evaluative impression of the defendant. Similarly, Payne, Lambert, and Jacoby (2002) reported that a specific warning failed to reduce the *weapons false alarm* effect (a higher rate of false perception of a weapon when

primed with a Black vs. White face), which is based on automatic stereotype activation, even though the warning did affect participants' self-reported goals.

We, too, found it impossible to entirely eliminate the CIE. Even when combining what are now known to be the two best strategies, misinformation lingered to a significant degree. So why is this effect so difficult to overcome? One possibility is that, as is the case in false memory studies, the impact of automatic retrieval processes (implicit activation, fluency, familiarity) may be very difficult to neutralize. "Inoculation by knowledge" (Gallo et al., 1997) will, thus, be successful, but only to the degree to which these powerful cognitive mechanisms can be overcome strategically, as is discussed in the previous section.

Real-World Implications

Our study has clear implications for real-world issues. Until now, the provision of an alternative account was the only known way to reliably reduce continued reliance on misinformation. The utility of this strategy is limited because when information must be corrected, an alternative account may not always be available. Warnings, by contrast, can provide a very powerful strategy to reduce people's reliance on misinformation.

The script we used in our experiments resembled a newspaper article. Notably, although retractions and corrections also appear in newspapers from one day to the next, the media abound with cases in which news items are qualified and corrected, even within a single report, as is the case when the title of an article does not exactly match its content. For example, the title of a front-page *New York Times* article in December 2008, "Dollar shift: Chinese pockets filled as Americans' emptied," (Landler, 2008) suggested that China had directly profited from the U.S. economy's crisis. This was corrected in a subsequently published "postscript" (and the title of the online article changed) to better match the topic of the article—namely, that U.S. consumption had largely relied on money borrowed from other countries, including China.⁶ Our study shows that specific information about the existence of the CIE, which could be included in media education, can help inoculate people against excessive reliance on misinformation.

In judicial settings, warnings seem to be a promising and economic way to reduce the rate at which jurors continue to rely on outdated or inadmissible information. Specific up-front warnings could be given to jurors routinely to build the expectation that they may face information in the media or in the actual trial that they will later have to disregard. Importantly, as our fact-recall data have again shown, such warnings will specifically decrease reliance on misinformation without having an impact on memory for veridical information.

CONCLUSION

Warnings have been found to be effective in reducing unwanted cognitive effects including false memories (Gallo et al., 1997; McDermott & Roediger, 1998), the impact of postevent suggestions (Echterhoff et al., 2005; Greene et al.,

1982) and misleading questions (Warren, Hulse-Trotter, & Tubbs, 1991), acquiescent responding in children (Saywitz & Moan-Hardie, 1994), imagination inflation effects on recognition confidence (Landau & von Glahn, 2004), and even cross-race effects in face recognition (Hugenberg, Miller, & Claypool, 2007). The present study found that informing people about an unwanted, but prevalent, cognitive effect—the CIE of misinformation—can likewise help them reduce or even avoid this processing error.

AUTHOR NOTE

Preparation of the present article was facilitated by a Discovery Grant and an Australian Professorial Fellowship from the Australian Research Council to S.L. We thank Charles Hanich, Kate Stephanou, Briony Swire, and Abel Oh for research assistance. Address correspondence to U. K. H. Ecker or S. Lewandowsky, School of Psychology (M304), University of Western Australia, Crawley, WA 6009, Australia (e-mail: ullrich.ecker@uwa.edu.au or lewan@psy.uwa.edu.au; Web site: www.cogsciwa.com).

REFERENCES

- ANASTASI, J. S., RHODES, M. G., & BURNS, M. C. (2000). Distinguishing between memory illusions and actual memories using phenomenological measurements and explicit warnings. *American Journal of Psychology*, *113*, 1-26.
- AYERS, M. S., & REDER, L. M. (1998). A theoretical review of the misinformation effect: Predictions from an activation-based memory model. *Psychonomic Bulletin & Review*, *5*, 1-21.
- 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.
- BUSH, J. G., JOHNSON, H. M., & SEIFERT, C. M. (1994). The implications of corrections: Then why did you mention it? In A. Ram & K. Eiselt (Eds.), *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society* (pp. 112-117). Hillsdale, NJ: Erlbaum.
- 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.
- 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.
- DEESE, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*, 17-22.
- EAKIN, D. K., SCHREIBER, T. A., & SERGENT-MARSHALL, S. (2003). Misinformation effects in eyewitness memory: The presence and absence of memory impairment as a function of warning and misinformation accessibility. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *29*, 813-825.
- ECHTERHOFF, G., HIRST, W., & HUSSY, W. (2005). How eyewitnesses resist misinformation: Social postwarnings and the monitoring of memory characteristics. *Memory & Cognition*, *33*, 770-782.
- ECKER, U. K. H., LEWANDOWSKY, S., & APAL, J. (in press). Terrorists brought down the plane!—No, actually it was a technical fault: Processing corrections of emotive information. *Quarterly Journal of Experimental Psychology*.
- ECKER, U. K. H., LEWANDOWSKY, S., SWIRE, B., & CHANG, D. (2010). *Misinformation in memory: Effects of the encoding strength and strength of retraction*. Manuscript submitted for publication.
- FEIN, S., MCCLOSKEY, A. L., & TOMLINSON, T. M. (1997). Can the jury disregard that information? The use of suspicion to reduce the prejudicial effects of pretrial publicity and inadmissible testimony. *Personality & Social Psychology Bulletin*, *23*, 1215-1226.
- GALLO, D. A., ROBERTS, M. J., & SEAMON, J. G. (1997). Remembering words not presented in lists: Can we avoid creating false memories? *Psychonomic Bulletin & Review*, *4*, 271-276.
- GALLO, D. A., ROEDIGER, H. L., III, & McDERMOTT, K. B. (2001). Associative false recognition occurs without strategic criterion shifts. *Psychonomic Bulletin & Review*, *8*, 579-586.
- GILBERT, D. T., KRULL, D. S., & MALONE, P. S. (1990). Unbelieving the unbelievable: Some problems in the rejection of false information. *Journal of Personality & Social Psychology*, *59*, 601-613.
- GILBERT, D. T., TAFARODI, R. W., & MALONE, P. S. (1993). You can't not believe everything you read. *Journal of Personality & Social Psychology*, *65*, 221-233.
- GOLDSMITH, M., KORIAT, A., & WEINBERG-ELIEZER, A. (2002). Strategic regulation of grain size in memory reporting. *Journal of Experimental Psychology: General*, *131*, 73-95.
- GREENE, E., FLYNN, M. S., & LOFTUS, E. F. (1982). Inducing resistance to misleading information. *Journal of Verbal Learning & Verbal Behavior*, *21*, 207-219.
- HUGENBERG, K., MILLER, J., & CLAYPOOL, H. M. (2007). Categorization and individuation in the cross-race recognition deficit: Toward a solution to an insidious problem. *Journal of Experimental Social Psychology*, *43*, 334-340.
- JOHNSON, H. M., & SEIFERT, C. M. (1994). Sources of the continued influence effect: When misinformation in memory affects later inferences. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *20*, 1420-1436.
- JOHNSON, H. M., & SEIFERT, C. M. (1998). Updating accounts following a correction of misinformation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *24*, 1483-1494.
- JOHNSON, H. M., & SEIFERT, C. M. (1999). Modifying mental representations: Comprehending corrections. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 303-318). Mahwah, NJ: Erlbaum.
- JOHNSON, M. K., HASHTRUDI, S., & LINDSAY, S. D. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3-28.
- JOU, J., & FOREMAN, J. (2007). Transfer of learning in avoiding false memory: The roles of warning, immediate feedback, and incentive. *Quarterly Journal of Experimental Psychology*, *60*, 877-896.
- KASSIN, S. M., & SUKEL, H. (1997). Coerced confessions and the jury: An experimental test of the "harmless error" rule. *Law & Human Behavior*, *21*, 27-46.
- KINTSCH, W., & GREENE, E. (1978). The role of culture-specific schemata in the comprehension and recall of stories. *Discourse Processes*, *1*, 1-13.
- KULL, S., RAMSAY, C., & LEWIS, E. (2003). Misperceptions, the media, and the Iraq War. *Political Science Quarterly*, *118*, 569-598.
- LANDAU, J. D., & VON GLAHN, N. (2004). Warnings reduce the magnitude of the imagination inflation effect. *American Journal of Psychology*, *117*, 579-593.
- LANDLER, M. (2008, December 26). Dollar shift: Chinese pockets filled as Americans' emptied. *The New York Times*. Online version, titled "Chinese savings helped inflate American bubble," retrieved November 11, 2009, from www.nytimes.com/2008/12/26/world/asia/26addiction.html?_r=1
- LEWANDOWSKY, S., STRITZKE, W. G. K., OBERAUER, K., & MORALES, M. (2005). Memory for fact, fiction and misinformation: The Iraq War 2003. *Psychological Science*, *16*, 190-195.
- LEWANDOWSKY, S., STRITZKE, W. G. K., OBERAUER, K., & MORALES, M. (2009). Misinformation and the "War on Terror": When memory turns fiction into fact. In W. G. K. Stritzke, S. Lewandowsky, D. Denemark, J. Clare, & F. Morgan (Eds.), *Terrorism and torture: An interdisciplinary perspective* (pp. 179-203). Cambridge: Cambridge University Press.
- LINDSAY, D. S., & JOHNSON, M. K. (1989). The eyewitness suggestibility effect and memory for source. *Memory & Cognition*, *17*, 349-358.
- LOFTUS, E. F. (1975). Leading questions and the eyewitness report. *Cognitive Psychology*, *7*, 560-572.
- LOFTUS, E. F., & PALMER, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning & Verbal Behavior*, *13*, 585-589.
- MARINI, A., BOEWE, A., CALTAGIRONE, C., & CARLOMAGNO, S. (2005). Age-related differences in the production of textual descriptions. *Journal of Psycholinguistic Research*, *34*, 439-463.
- MCCABE, D. P., & SMITH, A. D. (2002). The effect of warnings on false memories in young and older adults. *Memory & Cognition*, *30*, 1065-1077.

- McCloskey, M., & Zaragoza, M. (1985). Misleading postevent information and memory for events: Arguments and evidence against memory impairment hypotheses. *Journal of Experimental Psychology: General*, **114**, 1-16.
- McDERMOTT, K. B. (1996). The persistence of false memories in list recall. *Journal of Memory & Language*, **35**, 212-230.
- McDERMOTT, K. B., & ROEDIGER, H. L., III (1998). Attempting to avoid illusory memories: Robust false recognition of associates persists under conditions of explicit warnings and immediate testing. *Journal of Memory & Language*, **39**, 508-520.
- NEUSCHATZ, J. S., BENOIT, D. G., & PAYNE, D. G. (2003). Effective warnings in the Deese/Roediger-McDermott false-memory paradigm: The role of identifiability. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **29**, 35-41.
- NEUSCHATZ, J. S., PAYNE, D. G., LAMPINEN, J. M., & TOGLIA, M. P. (2001). Assessing the effectiveness of warnings and the phenomenological characteristics of false memories. *Memory*, **9**, 53-71.
- PAYNE, B. K., LAMBERT, A. J., & JACOBY, L. L. (2002). Best laid plans: Effects of goals on accessibility bias and cognitive control in race-based misperceptions of weapons. *Journal of Experimental Social Psychology*, **38**, 384-396.
- PETERS, M. J. V., JELICIC, M., GORSKI, B., SIJSTERMANS, K., GIESBRECHT, T., & MERCKELBACH, H. (2008). The corrective effects of warning on false memories in the DRM paradigm are limited to full attention conditions. *Acta Psychologica*, **129**, 308-314.
- RANTZEN, A., & MARKHAM, R. (1992). The reversed eyewitness testimony design: More evidence for source monitoring. *Journal of General Psychology*, **119**, 37-43.
- RAPP, D. N., & KENDEOU, P. (2007). Revising what readers know: Updating text representations during narrative comprehension. *Memory & Cognition*, **35**, 2019-2032.
- ROEDIGER, H. L., III, & McDERMOTT, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 803-814.
- ROSENTHAL, R., ROSNOW, R. L., & RUBIN, D. B. (2000). *Contrasts and effect sizes in behavioral research: A correlational approach*. Cambridge: Cambridge University Press.
- SAYWITZ, K. J., & MOAN-HARDIE, S. (1994). Reducing the potential for distortion of childhood memories. *Consciousness & Cognition*, **3**, 408-425.
- SCHUL, Y., & MANZURY, F. (1990). The effect of type of encoding and strength of discounting appeal on the success of ignoring an invalid testimony. *European Journal of Social Psychology*, **20**, 337-349.
- SCHWARZ, N., SANNA, L. J., SKURNIK, I., & YOON, C. (2007). Metacognitive experiences and the intricacies of setting people straight: Implications for debiasing and public information campaigns. *Advances in Experimental Social Psychology*, **39**, 127-161.
- SEAMON, J. G., LUO, C. R., KOPECKY, J. J., PRICE, C. A., ROTHSCHILD, L., FUNG, N. S., & SCHWARTZ, M. A. (2002). Are false memories more difficult to forget than accurate memories? The effect of retention interval on recall and recognition. *Memory & Cognition*, **30**, 1054-1064.
- SEIFERT, C. M. (2002). The continued influence of misinformation in memory: What makes a correction effective? In B. H. Ross (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 41, pp. 265-292). San Diego: Academic Press.
- VAN OOSTENDORP, H. (1996). Updating situation models derived from newspaper articles. *Medienpsychologie*, **8**, 21-33.
- VAN OOSTENDORP, H., & BONEBAKKER, C. (1999). Difficulties in updating mental representations during reading news reports. In H. van Oostendorp & S. R. Goldman (Eds.), *The construction of mental representations during reading* (pp. 319-339). Hillsdale, NJ: Erlbaum.
- WARREN, A., HULSE-TROTTER, K., & TUBBS, E. C. (1991). Inducing resistance to suggestibility in children. *Law & Human Behavior*, **15**, 273-285.
- WATSON, J. M., McDERMOTT, K. B., & BALOTA, D. A. (2004). Attempting to avoid false memories in the Deese/Roediger-McDermott paradigm: Assessing the combined influence of practice and warnings in young and old adults. *Memory & Cognition*, **32**, 135-141.
- WESTERBERG, C. E., & MARSOLEK, C. J. (2006). Do instructional warnings reduce false recognition? *Applied Cognitive Psychology*, **20**, 97-114.
- WHITTLESEA, B. W. A., & WILLIAMS, L. D. (2001). The discrepancy-attribution hypothesis: I. The heuristic basis of feelings of familiarity. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 3-13.
- WILKES, A. L., & LEATHERBARROW, M. (1988). Editing episodic memory following the identification of error. *Quarterly Journal of Experimental Psychology*, **40A**, 361-387.
- WILKES, A. L., & REYNOLDS, D. J. (1999). On certain limitations accompanying readers' interpretations of corrections in episodic text. *Quarterly Journal of Experimental Psychology*, **52A**, 165-183.
- WILSON, T. D., & BREKKE, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin*, **116**, 117-142.
- ZARAGOZA, M. S., & KOSHMIDER, J. W. (1989). Misled subjects may know more than their performance implies. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 246-255.
- ZARAGOZA, M. S., & LANE, S. M. (1994). Source misattributions and the suggestibility of eyewitness memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 934-945.

NOTES

1. We use the label *misinformation* for a piece of information that is corrected after its initial dissemination, without meaning to imply that it was initially introduced in order to intentionally mislead the receiver.
2. The last recall question that directly queried who was on the bus was not included in the fact-recall score because there was no unambiguously correct response to this question in most conditions (this question did contribute to the inference score, however). Thus, the maximum recall score was 8.
3. This is not always the case (e.g., Ecker et al., in press), so the establishment of the true baseline was considered to be an important control factor.
4. This pattern of results was also obtained in the less powerful analysis on the initial smaller sample, with $n = 25$ in the specific warning + alternative condition and with $n = 22$ in the corresponding restricted subsample (participants who remembered the retraction).
5. Some of the explanations put forward, however, do not apply to the current paradigm. For example, it has been suggested that warnings lead to shallower encoding of misinformation (Echterhoff et al., 2005). Participants in a continued influence paradigm cannot strategically encode misinformation less deeply, because, at the time of encoding, the misinformation status is not evident.
6. We thank Craig Silverman (www.craigsilverman.ca) for pointing out this example.

APPENDIX A
Account of Minibus Accident

Messages Used in Experiment 1 (Critical Messages Italicized)

Message 1. A report has come in to the police headquarters on Sunday about a serious accident involving a minibus. The report came from a driver who was driving past the scene of the accident.

Message 2. The minibus had crashed into a steep embankment near Spring Street and had rolled on its side. This resulted in injuries to some of the passengers on board.

Message 3. A rescue crew was dispatched to the scene immediately upon report of the accident. They arrived at the scene within ten minutes.

Message 4 [misinformation]. Police have stated that the passengers on the bus were a group of elderly people who were on their way back to their nursing home after a bingo game. The weather was reportedly fine and visibility was good. No other vehicles seem to have been involved.

Message 5. The rescue crew started evacuation of the minibus and wanted to reach the injured passengers first but found it difficult to tell them apart from the uninjured.

Message 6. The rescue crew also reported difficulty in getting both injured and uninjured passengers out of the minibus even though the exits were clear.

Message 7. Authorities are desperately trying to reach the family members of the victims involved in the accident and managed to trace the minibus license number to a rental company.

Message 8. The rescue crew reported that the rescue efforts were slow and would take more time and advised bystanders not to crowd around the area. Due to building-up traffic, the local radio station advised drivers to avoid the Spring Street area.

Message 9. Local television shows live footage of uninjured passengers having problems getting up the embankment.

Message 10. Rescue crew can be heard remarking that the uninjured passengers were unable to help in the rescue efforts.

Message 11 [no-retraction control condition]. A second statement from the police has confirmed that all passengers were successfully rescued.

Message 11 [retraction-only, general warning, and specific warning conditions]. A second statement from the police stated that the passengers on the minibus were not elderly people.

Message 11 [alternative condition]. A second statement from the police has stated that the passengers on the minibus were not elderly people but college hockey players returning from a victory party after the state hockey final.

Message 12. Some passengers with injuries were taken to the nearby St Joseph's hospital for treatment.

Message 13. At the hospital, 3 passengers with more serious injuries had to be warded for observation while the others were discharged after treatment.

Message 14. Some of the uninjured passengers interviewed at the scene of the accident remarked how helpless and dependent they were on the rescue crew during the accident and expressed their gratitude to them.

Alternative Message Used in Alternative-Throughout Condition of Experiment 2

Message 4. Police have stated that the passengers on the bus were college hockey players returning from a victory party after the state hockey final. The weather was reportedly fine and visibility was good. No other vehicles seem to have been involved.

APPENDIX B
Warnings

General Warning

In their desire to sensationalize, the media sometimes does not check facts before publishing information that turns out to be inaccurate. It is therefore important to read the following story and answer the questions at the end carefully.

Specific Warning

Research has shown that people continue to rely on outdated information even when it has been retracted or corrected. An example of this is the fact that some people today still believe that Iraq had weapons of mass destruction even though none have been found. Another example is court proceedings in which jurors continue to use information that has been deemed inadmissible. Please read the following story carefully and answer the questions at the end without relying on any outdated information.

APPENDIX C
Questionnaire

Inference Questions

1. Why do you think it was hard telling apart who the injured and uninjured passengers were?
2. Why do you think it was difficult getting both the injured and uninjured passengers out of the minibus?
3. Which family members of passengers are authorities most likely to contact to inform them about the accident?
4. Why do you think it was difficult getting the uninjured passengers up the embankment?
5. Why do you think the uninjured passengers were unable to help with the rescue efforts?
6. Why do you think some passengers were injured while others were not?
7. How could such an incident be avoided in the future?
8. Why did the uninjured passengers feel helpless and dependent on the rescue crew?
9. Why do you think the minibus crashed?

Fact Questions

1. On which day did the accident occur?
2. Who reported the accident to the police?
3. Where did the minibus land after the accident?
4. Where did the accident occur?
5. How was the weather like on that day?
6. Which hospital were the injured taken to?
7. How many injured passengers were warded in the hospital?
8. Where were the uninjured passengers interviewed?
9. Who was on the minibus during the accident?

Manipulation-Check Questions

1. What was the purpose of the second message from the police?
2. Were you aware of any corrections in the messages that you read?

Debriefing Questions (Used Only in Extension of Experiment 2)

1. What do you think was the purpose of this experiment?
 2. What was the false information given in the story?
 3. What was the corrected information given in the story?
 4. Did you believe the corrected version of the event?
-

(Manuscript received November 13, 2009;
revision accepted for publication May 15, 2010.)