The Quarterly Journal of Experimental Psychology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t716100704

Terrorists brought down the plane!—No, actually it was a technical fault: Processing corrections of emotive information

Ullrich K. H. Ecker[a]; Stephan Lewandowsky[a]; Joe Apai[a]
[a] University of Western Australia, Crawley, WA, Australia

First published on: 06 August 2010

To cite this Article Ecker, Ullrich K. H. , Lewandowsky, Stephan and Apai, Joe(2011) 'Terrorists brought down the plane!—No, actually it was a technical fault: Processing corrections of emotive information', The Quarterly Journal of Experimental Psychology, 64: 2, 283 — 310, First published on: 06 August 2010 (iFirst)
To link to this Article: DOI: 10.1080/17470218.2010.497927
URL: http://dx.doi.org/10.1080/17470218.2010.497927

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2011, 64 (2), 283–310

Psychology Press
Taylor & Francis Group

Terrorists brought down the plane!—No, actually it was a technical fault: Processing corrections of emotive information

Ullrich K. H. Ecker, Stephan Lewandowsky, and Joe Apai
University of Western Australia, Crawley, WA, Australia

It is well known that people often continue to rely on initial misinformation even if this information is later corrected and even if the correction itself is remembered. This article investigated the impact of emotionality of the material on people's ability to discount corrected misinformation. The focus was on moderate levels of emotionality comparable to those elicited by real-world news reports. Emotionality has frequently been shown to have an impact upon reasoning and memory, but the generality of this influence remains unclear. In three experiments, participants read a report of a fictitious plane crash that was initially associated with either an emotionally laden cause (terrorist attack) or an emotionally more neutral cause (bad weather). This initial attribution was followed by a retraction and presentation of an alternative cause (faulty fuel tank). The scenarios demonstrably affected participants' self-reported feelings. However, all three experiments showed that emotionality does not affect the continued influence of misinformation.

Keywords: Continued influence of misinformation; Memory; Emotionality.

In today's information society, new pieces of information frequently demand changes to already-memorized event models—new information needs to be integrated, and old information needs to be revised, updated, or discarded. For example, at least three successive theories concerning responsibility for the 2001 anthrax attacks in the United States have been publicly put forward, ranging from Saddam Hussein's Iraq to first one and then another American weapons scientist. Each theory was presented as factual—or was at least strongly intimated—by authorities before being replaced by a new official version or being called into question by media reports. Such changes to event models require a sophisticated revision process.

The continued influence of misinformation

Previous research on the comprehension of ongoing, unfolding events has shown that this revision process is error prone. In particular,

Correspondence should be addressed to Ullrich K. H. Ecker, School of Psychology, University of Western Australia, Crawley, WA 6009, Australia. E-mail: ullrich.ecker@uwa.edu.au

Preparation of this paper was facilitated by a Discovery Grant and an Australian Professorial Fellowship from the Australian Research Council to Stephan Lewandowsky. The lab's website is located at http://www.cogsciwa.com. We thank Klaus Oberauer and Jeff Rouder for comments on an earlier version, E.-J. Wagenmakers and Ruud Wetzels for advice on the Bayesian analysis, and Charles Hanich, Kate Stephanou, Stewart Craig, David Tang, Briony Swire, and Laura Cook for research assistance.

http://www.psypress.com/qjep
DOI:10.1080/17470218.2010.497927

people often continue to rely on information even after the initial information has been invalidated by further evidence and has been reported to be false (de Vega, Urrutia, & Riffo, 2007; H. M. Johnson & Seifert, 1994; Wilkes & Leatherbarrow, 1988). This *continued influence of misinformation*[1] arises even if participants demonstrably understand and remember a subsequent retraction or correction. The predominant explanation of this effect is that people prefer to have an *incorrect* event-model to having an *incomplete* event-model; that is, if the retraction of a central event aspect (such as the cause of an event) leaves a significant gap in the event-model, people tend to rely on outdated information even if they know it has been discredited, rather than acknowledging the void (Ecker, Lewandowsky, & Tang, in press; H. M. Johnson & Seifert, 1994; van Oostendorp, 1996; van Oostendorp & Bonebakker, 1999; Wilkes & Leatherbarrow, 1988).

Notably, this pattern of continued influence has not been confined to arbitrary stimuli but has also been observed in situations of real-world significance. For example, (mock) jurors often rely on evidence that has been designated inadmissible, despite claiming to ignore it (Fein, McCloskey, & Tomlinson, 1997; Kassin & Sukel, 1997). Likewise, public information campaigns are potentially ineffective or even counterproductive if they list "facts and myths about . . . ", because attempts to explicitly discredit false information necessarily involve the repetition of these myths, which in itself strongly reinforces belief in them (Schwarz, Sanna, Skurnik, & Yoon, 2007). Finally, investigations of people's understanding of events surrounding the US-led invasion of Iraq in 2003 have repeatedly revealed that substantial segments of the American public continued to believe that weapons of mass destruction (WMD) had been found in Iraq after the US-led invasion (Kull, Ramsay, & Lewis, 2003; Lewandowsky, Stritzke, Oberauer, & Morales, 2005, 2009; Nyhan & Reifler, 2010).

Notwithstanding the prominence of continued influence of misinformation, the effect is not ubiquitous, and a number of boundary conditions have been identified: First, it only seems to occur if the information is initially of high causal relevance for a focal event or outcome (H. M. Johnson & Seifert, 1994; van Oostendorp, 1996). Second, and most important, the effect is diminished or abolished if the correction involves more than a simple negation—for example, if it also provides a causal alternative that can replace the role originally held by the initial misinformation and hence fill the event-model gap left behind by the retraction (H. M. Johnson & Seifert, 1994; Rapp & Kendeou, 2007). To illustrate, if a person is initially suspected of a crime, people's belief in that person's guilt may persist even when an alibi is presented, but it will be significantly reduced if the alibi is accompanied by presentation of an alternative suspect. Finally, along related lines, Fein et al. (1997) and Lewandowsky et al. (2005, 2009) proposed that people's readiness to discount initial misinformation increases if they are *suspicious* of the motives underlying its dissemination; either that suspicion can be induced experimentally (e.g., by informing participants that they may be deliberately misled; Fein et al., 1997) or it can be the result of preexisting scepticism[2] (e.g., Lewandowsky et al., 2005, found that people who questioned the official WMD-related reasons for the invasion of Iraq were more likely to discount war-related news reports that later turned out to be false than people who accepted that the war was fought over WMDs). Suspicion and general scepticism may therefore act as mechanisms that facilitate the endogenous generation or consideration of potential causal alternatives.

[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 concept of scepticism refers to a more stable personality trait that facilitates the emergence of specific suspicions and can be measured independently of the misinformation issue under investigation (cf. Lewandowsky et al., 2009).

What remains unknown, however, is whether those boundary conditions also apply to information of negative emotional valence: Have people discounted initial suspicions of Iraqi involvement in the 2001 anthrax attacks even though not just one but two alternative suspects have been presented? In today's world, would people relinquish initial suspicions of Islamic involvement after a terrorist incident even when the perpetrator turns out to be a former U.S. serviceman?—as was the case in the Oklahoma City bombing, in which initial suspicion fell onto a Muslim citizen before Timothy McVeigh was arrested.

The effects of emotion on memory and memory updating

Emotion can affect memory updating at two distinct stages: (a) It may affect the encoding of the initial information (irrespective of whether or not it is later retracted); and (b) it may affect people's ability to update or alter that information.

It is often claimed that emotional information leads to better memory (e.g., Levine & Pizarro, 2004). This has been demonstrated numerous times for autobiographical material (e.g., Conway et al., 1994; Thompson, Skowronski, Larsen, & Betz, 1996), including traumatic experiences that result in highly vivid "flashbulb memories" (Brown & Kulik, 1977; Christianson & Loftus, 1990), but it also arises with moderately emotive material in the laboratory. For example, Heuer and Reisberg (1990) and Burke, Heuer, and Reisberg (1992) reported that central aspects or the "gist" of a story were better recalled if the story was emotional. Likewise, aversive pictures (e.g., Hamann, Ely, Grafton, & Kilts, 1999) and emotive words (e.g., Kensinger, 2008; Kensinger & Corkin, 2003a) have been found to yield better memory than more neutral materials. Inasmuch as emotional material is better remembered, it follows that initial misinformation that is high in emotionality may be more resistant to subsequent correction or retraction than less emotional information.

However, the effects of emotion on memory can also be detrimental. For example, information that is not central to an emotional event can be negatively affected by emotionality (Burke et al., 1992; Christianson & Loftus, 1987, 1991; Pezdek, 2003; Schmidt, 2004; see Heuer & Reisberg, 1990, for a null effect). Even flashbulb memories have been shown to be prone to reconstructive errors despite the very high levels of vividness and confidence with which they are held (Christianson, 1989; McCloskey, Wible, & Cohen, 1988; Neisser & Harsch, 1992; Schmidt, 2004). On balance, therefore, it appears that the facilitative effects of emotion on memory are primarily limited to the retrieval of schematic highlights (e.g., that the event actually happened) but are largely absent—or indeed harmful—for more peripheral details.

Similar ambivalence surrounds the likely role of emotion during updating or retraction of information. Although to our knowledge the effect of emotionality on the correction of initial misinformation has not yet been tested, the literature does offer some hints: On the one hand, emotional information may result in a heightened level of arousal which in itself may reduce subsequently available processing resources (e.g., by blocking working memory; cf. Eysenck & Calvo, 1992; Ridley & Clifford, 2004). It follows that emotive information may be harder to correct than neutral material. A converging line of research originates with the source monitoring framework proposed by M. K. Johnson, Hashtroudi, and Lindsay (1993). Source monitoring refers to the ability to remember *where* the information originated; for example, whether it was read or heard or relayed by a friend. It has been suggested that there is a source monitoring advantage for emotional stimuli relative to nonemotional ones, implying that more contextual, source-defining features can be retrieved for emotional stimuli (D'Argembeau & van der Linden, 2004; Doerksen & Shimamura, 2001; Kensinger & Schacter, 2006; MacKay et al., 2004). It follows that memory for emotive information might be more stable and less prone to distortion and thus perhaps also less prone to correction. The assumption that emotionality—in particular of negative valence—leads to better and more stable memory

representations of central stimulus aspects seems to be the dominant view in the current literature (cf. Kensinger, 2009).

Opposing these lines of argument, however, is other recent research that can be taken to suggest that emotional material may in fact be more susceptible to memory distortions than neutral material in at least some circumstances. For example, negative emotional valence can drive false memories because it seems to lead to less specific item memory (i.e., lower verbatim recollection but higher gist familiarity for actually studied items; Brainerd, Stein, Silveira, Rohenkohl, & Reyna, 2008; see also Porter, Taylor, & ten Brinke, 2008). Such familiarity can even lead to flashbulb-like memories for events that were never witnessed (e.g., remembering seeing a particular plane crash on TV although there was no video footage of the crash; Crombag, Wagenaar, & van Koppen, 1996; Ost, Granhag, Udell, & af Hjelmsäter, 2008). There is also evidence that emotionality may impair contextual integration (e.g., the binding of an item to its source context; Mather, 2007; Mather et al., 2006). It follows on this alternative line of reasoning that the representation of emotional information may be more susceptible to subsequent manipulation because participants may mistake parts of subsequent information for aspects of the initial information (i.e., source confusion; cf. Ridley & Clifford, 2004, for a similar argument). In cases where later information corrects or retracts earlier information, this source confusion—which normally has adverse consequences—may translate into an overall beneficial effect on memorial accuracy.

The present study

In summary, despite the dominant view that emotionality leads to more stable memory representations of central stimulus aspects, previous research leads to somewhat inconsistent predictions about the impact of emotion on the discounting of initial misinformation.

The aim of the present study was to investigate the influence of initial misinformation in moderately negative emotive material, of the kind we encounter every day in newspaper reports and TV news. Healy, Aylward, Bourne, and Beer (2009) demonstrated that even clearly fictitious news reports of repeated terrorist attacks can lead to an escalation in self-reported emotions such as fear and anger. Unz, Schwab, and Winterhoff-Spurk (2008) investigated emotional responses to real TV news and found that reports of intentional violence (e.g., assaults, terrorist attacks) in particular led to increased levels of negative emotions such as anger, sadness, and disgust, relative to reports of unintentional violence such as accidents. The question we ask is: How does the emotionality associated with the encoding of such reports—which many people are exposed to regularly—influence the ability to discount (mis)information that later turns out to be false?

To this end, we now present three experiments that tested the effect of negative emotionality on the continued influence of misinformation by contrasting two initially presented causes of a fictitious plane crash: a terrorist attack (i.e., intentional violence) and bad weather (i.e., unintentional violence), with the former shown to be more emotive. In some conditions, the initially stated cause was retracted later in the scenario. In some of these cases the retraction was additionally accompanied by presentation of a more neutral alternative cause—namely, a faulty fuel tank. The provision of a causal alternative is known to enable people to discount initial misinformation with scenarios of little emotionality; is discounting also possible when the initial misinformation has high emotionality?

To foreshadow, we first present the three studies with the conventional null-hypothesis-testing analyses. Because the results converge on the conclusion that emotionality has no effect on how people process retractions, we follow up the conventional analyses with a Bayesian analysis, which permits a more forceful affirmation of null effects. This Bayesian analysis is presented in the General Discussion, once all experiments have been reported.

EXPERIMENT 1

The experiment employed a 2 × 3 between-subjects design. Conditions were defined by the orthogonal combination of emotionality of the initially presumed crash cause (two levels) and the type of retraction (three levels); there was an additional control group not involving any correction that did not form part of this design. Hence there were seven different versions of the plane crash scenario, each administered to a different group of participants.

Method

Participants

A total of 70 first-year psychology students participated for partial course credit. Participants were tested individually and were randomly assigned to the seven experimental conditions ($N = 10$ in each cell).

Stimuli

Participants were presented with a folder, containing a series of seven messages, each printed on a separate sheet of paper, which provided a fictitious account of a plane crash. (Materials of all experiments are provided in the Appendix.) All versions of the scenario were concerned with a plane crashing in New South Wales, Australia, en route from Sydney to Brisbane, killing over 100 passengers. The two design variables were implemented as follows.

First, the two levels of emotionality were instantiated by scenarios that differed with respect to the initially reported cause of the crash. The fourth message in the sequence contained the statement: "media . . . reported that . . . " either "bad weather . . . " (low emotionality) or "terrorists . . . " (high emotionality) "were to blame for the crash."

To check whether these two scenarios differed in likely emotionality, we performed a generic newspaper database search including Asian, Australian, UK, and U.S. sources at www.factiva.com. On the assumption that the psychological similarity between words is reflected by their co-occurrence in natural language (Landauer & Dumais, 1997; see also Bullinaria & Levy, 2007; Griffiths, Steyvers, & Firl, 2007), we searched for the terms "plane crash" AND (bomb OR terrorist OR terrorism) versus "plane crash" AND (weather OR thunderstorm), combined with any of the emotional words "fear", "scared", "anger", "angry", "sad", or "sadness". The base rates for "bomb OR terrorist OR terrorism" and "weather OR thunderstorm" were comparable (6,143,754 vs. 6,238,354). The number of hits for the emotional and neutral scenarios combined with the emotional words was 5,897 versus 3,353. These values were baseline corrected using the ratio reflecting slightly higher level of occurrence for "weather OR thunderstorm", which was 1.015. The distribution of the baseline-corrected scores differed significantly from expectation under equality, with continuity-corrected $\chi^2(1) = 742.74$, $p < .001$. Experiments 2 and 3, which used similar materials, included a direct empirical validation of the differing emotionality of the two types of scenario.

Second, the orthogonal retraction variable was instantiated by three different variants of the scenarios: In the no-retraction conditions the initially presented cause was never corrected. In the retraction-only conditions, the sixth message in the sequence contained the retraction (e.g., " . . . it was concluded that bad weather was not to blame"). In the causal-alternative conditions, the same retraction was additionally accompanied by a causal alternative presented in the same message ("The actual cause was determined to be a faulty fuel tank").

Finally, the control group (not part of the orthogonal design) received a scenario referring to the alternative (the faulty fuel tank) throughout, without any prior misinformation (i.e., terrorism or bad weather) and without retraction. This condition was included for two reasons. First, because the faulty fuel tank was the alternative explanation in the causal-alternative conditions, the control condition allowed us to compare the number of references to this cause between when it was presented as the initial (and uncorrected) cause and

when it was the alternative cause (and correcting initial misinformation). Second, the control condition supplied a baseline measure of the "misinformation" causes—that is, how often people spontaneously refer to bad weather or a bomb attack as the cause of a plane crash without being given this information.

Procedure

Participants read the news-like messages at their own pace without backtracking. People then engaged in a 5-minute distractor task before responding to a questionnaire comprising nine inference questions.

Inference questions were designed to elicit inferences about aspects of the incident that were directly or indirectly linked to its cause (e.g., "Why were there so few survivors?" or "What would be a more plausible alternative cause for the crash than the media's report?").

Results

Coding of responses

Analysis focused on people's reliance on (a) initial misinformation and (b) the alternative information in response to the inference questions. All references needed to be causal in nature; each item containing one or more causal references was given a score of 1. If a response referred to both initial misinformation and the alternative information, both references were scored as 0.5. The highest possible individual score for both references to initial misinformation and the alternative information was therefore 9.

References to initial misinformation

Any reference to bad weather or a terrorist attack causing the plane crash was considered a reference to initial misinformation (which, strictly speaking, was never identified as misinformation in the no-retraction conditions). For instance, the response "Terrorism hysteria" to the inference question "Why does the media's account of the plane crash seem plausible enough to report?" was not scored as a reference to initial misinformation because it was not causal (i.e., score 0). In contrast, a statement such as "Because of the bomb fragments that were found" would receive a score of 1. The average misinformation scores are shown in Table 1. Note that there were no spontaneous references to (unpresented) misinformation in the alternative-throughout control group, and so there was no need for baseline correcting the scores of the six experimental conditions.

A two-way analysis of variance (ANOVA) on all conditions (except the control group) with factors emotionality and retraction revealed a main effect of retraction, $F(2, 54) = 171.90$, $\eta^2 = .86$, $p < .001$. There was no significant effect of emotionality, $F(1, 54) = 3.96$, $\eta^2 = .07$, $p > .05$. Most importantly theoretically, selective failure to retract emotive material would result in an interaction of both experimental variables, but the interaction was clearly nonsignificant, $F < 1$.

Follow-up tests (Tukey's HSD, honestly significant difference) showed that aggregating across the low and high emotionality groups, the

Table 1. *Mean number of references to initial misinformation in Experiment 1*

	Emotionality					
	Low		*High*		*Combined*	
Retraction condition	*M*	*SE*	*M*	*SE*	*M*	*SE*
No-retraction	5.20	0.25	6.00	0.26	5.60	0.20
Retraction-only	6.30	0.26	6.90	0.46	6.60	0.27
Causal-alternative	0.80	0.33	1.00	0.37	0.90	0.24

Note: Number of participants in each cell is $N = 10$, and maximum score is 9.

no-retraction and retraction-only conditions differed from the causal-alternative conditions (all ps $< .001$). Differences between emotionality groups were assessed with more powerful planned comparisons, which failed to reveal any significant differences between emotionality groups at any retraction level (all Fs < 2.97, all ps $> .09$). Finally, the rate of references to initial misinformation was larger than zero even in the causal-alternative conditions, $t(19) = 3.76$, $p = .001$.

References to the alternative information

Any *causal* reference to a faulty fuel tank (e.g., "plane ran out of fuel due to a faulty fuel tank") was considered a reference to the alternative information. This information was only provided in the causal-alternative conditions (viz., as the causal alternative) and in the alternative-throughout control condition (where it was the only cause given). In the no-retraction and retraction-only conditions this information was never provided. References to the alternative information are presented in Table 2. The results for the alternative-throughout control condition—which provide an upper bound on the number of references to the alternative information—are not shown in the table; the mean number of references in that condition was $M = 5.30$, $SE = 0.40$.

Not surprisingly, there were virtually no references to the alternative information in the conditions that never received it, and hence these were not included in further analyses. A one-way ANOVA comparing both causal-alternative conditions to the alternative-throughout control condition revealed no significant effect, $F(2, 27) = 1.13$. Likewise, the planned contrast between low and high emotionality was nonsignificant, $F < 1$.

Discussion

Experiment 1 replicated the finding of H. M. Johnson and Seifert (1994) that a correction on its own does not reliably enable people to discount initial misinformation. Instead, discounting was successful only when the correction was accompanied by a causal alternative.

In addition, Experiment 1 yielded the novel finding that emotionality of the material had no effect on the continued influence of misinformation. Emotionally laden information did not prove more resistant to correction than more neutral information. However, before this conclusion can be accepted, we must address three potential limitations of Experiment 1: One concern involves the high levels of reference to initial misinformation in the retraction-only conditions (the entries in Table 1 translate into .70–.78 of the maximum possible score), which may have prevented the detection of an effect of emotionality. A second, related concern involves the fact that Experiment 1 yielded a null effect with respect to the variable of greatest interest—namely, emotionality (or more precisely, the Emotionality × Retraction interaction). Third, it is unclear whether this null effect occurred because our manipulation of emotionality was insufficiently powerful. Although there was clear evidence that the media are more likely to rely on

Table 2. *Mean number of references to correct information in Experiment 1*

	Emotionality					
	Low		*High*		*Combined*	
Retraction condition	*M*	*SE*	*M*	*SE*	*M*	*SE*
No-retraction	0.10	0.10	0.10	0.10	0.10	0.07
Retraction-only	0		0		0	
Causal-alternative	4.50	0.56	4.30	0.52	4.40	0.37

Note: Number of participants in each cell is $N = 10$, and maximum score is 9.

emotional language in connection with terrorism than other causes of plane crashes, Experiment 1 did not include a behavioural measure of the effectiveness of our emotionality manipulation.

Experiment 2 addressed these concerns. First, we extended the retention interval to 40 minutes. This may facilitate the emergence of an emotionality effect by (a) avoiding ceiling effects and (b) creating time for consolidation processes to operate to which the effects of emotions on memory have been ascribed (cf. Burke et al., 1992; McGaugh, 2000; Sharot & Yonelinas, 2008). Second, to buttress the implications of a possible null result, we (a) increased sample size and hence power and (b) sought to enhance the impact of the emotionality manipulation by moving the crash site close to the city in which the experiment was carried out and by using a city of departure that either is or is not frequently mentioned in connection with terrorism. Third, we directly tested the impact of the emotionality manipulation by measuring how the different scenarios affected participants' self-reported feelings immediately upon reading. Finally, Experiment 2 also tested fact recall, including a direct question concerning the retraction.

EXPERIMENT 2

Experiment 2 used the same 2×3 between-subjects design augmented by the alternative-throughout control group that was used in Experiment 1.

Method

Participants

A total of 112 third-year psychology students participated in this experiment for partial course credit (17 males, mean age 21.4 years, age range 19–43 years). None of them had taken part in Experiment 1. An equal number of participants ($N = 16$) were randomly assigned to each of the seven conditions.

Stimuli

The stimuli were similar to those used in Experiment 1, except that the fictitious plane was headed for and crashed near Perth, Western Australia. In the low emotionality conditions, the plane departed in Auckland, New Zealand (going via Melbourne, Australia) before crashing due to bad weather. In the high emotionality conditions, the plane originated in Tehran, Iran (via Dubai, UAE) before crashing due to a terrorist bomb.[3]

To verify the impact of the emotionality manipulation, a separate sample of 32 participants (10 males, mean age 20.7 years, age range 18–36 years) read the no-retraction scenarios of either high or low emotionality. Participants then completed two visual analog scales (VAS) with poles happy–sad and angry–calm, respectively. Each VAS was a 10-cm horizontal line on which participants placed a mark to record their current emotional state. Participants who read the emotive scenario rated themselves as sadder ($M = 7.6$ cm, $SE = 0.4$) and angrier ($M = 5.1$ cm, $SE = 0.6$) than people who read the more neutral scenario (sadness: $M = 6.4$ cm; $SE = 0.4$; anger: $M = 3.8$ cm; $SE = 0.5$). One-way ANOVAs showed the difference to be significant for sadness, $F(1, 30) = 5.59$, $\eta^2 = .157$, $p < .03$, and (marginally) for anger, $F(1, 30) = 2.90$, $\eta^2 = .088$, $p < .10$.

Participants then completed the Positive Affect and Negative Affect Schedule (PANAS; Watson, Clark, & Tellegen, 1988) based on how the story made them feel. The PANAS assesses positive

[3] The association of terrorism and Islam has become an increasingly common stereotype in Western countries (Cainkar, 2002; Sheridan, 2006). Thus, choosing an Islamic city of departure was intended to suggest a religiously motivated terrorist attack and hence make the story more emotionally charged (cf. Hayes, Schimel, & Williams, 2008; Solomon, Greenberg, & Pyszczynski, 1991). A database search on factiva.com yielded 14,599 hits searching for "(terrorism OR terrorist) AND bomb AND Tehran" but only 1,393 when replacing Tehran with Auckland (the base rate of hits for "Tehran" and "Auckland" alone was 628,298 and 1,089,924, respectively). The distribution of the baseline-corrected scores differed significantly from an equal distribution, with continuity-corrected $\chi^2(1) = 21{,}346$, $p < .001$.

and negative affect with 10 items each (e.g., "upset" or "excited"; rated on a 5-point Likert scale). Participants who read the emotive scenario scored higher on the negative affect scale ($M = 17.50$, $SE = 1.64$) than people who read the low-emotionality scenario ($M = 14.19$; $SE = 0.84$), but there was no difference on the positive affect scale (high emotionality: $M = 16.69$, $SE = 1.18$; low emotionality: $M = 16.38$; $SE = 0.81$). An ANOVA found the differences of the difference scores—positive minus negative affect—to be significant, $F(1, 30) = 4.18$, $\eta^2 = .122$, $p = .05$.

We conclude that our scenarios successfully instantiated the desired emotionality manipulation.

Procedure

The procedure was identical to that of Experiment 1 with the following exceptions. First, the retention interval was extended to approximately 40 minutes, during which participants participated in an unrelated experiment involving several working-memory tasks. Second, the inference questionnaire (which now contained 12 questions) was followed by 12 fact recall questions that did not directly relate to the cause of the plane crash (e.g., "Where did the plane crash occur?"). One of the questions targeted recall of the retraction ("Was this report challenged in any way?").

Results

Coding of responses

Responses were scored by a trained scorer who was blind to condition. A second, independent judge scored 35 (5 per condition) randomly selected scoring sheets to calculate an inter-rater reliability index for causal references, which was high ($r = .93$).

Each fact recall question was scored with either 1 (correct) or 0 (incorrect). The question directly targeting the retraction—which was analysed separately—and a further question that could not be unambiguously scored in every condition were not included in the recall score, so the maximum score was 10.

The inference questions were scored as in Experiment 1, except that cause-related inferences that were contained in fact-recall responses also contributed to the inference scores (for instance, in response to the recall question "How many survivors were there?", a participant may have noted that "All passengers were killed by the bomb"; although this would be an incorrect response in terms of fact recall—there were survivors—it would be scored as an inference to initial misinformation). The highest possible score for references to initial misinformation and the alternative information was thus 24.

Fact recall

A 2×3 ANOVA on fact recall revealed no significant effects (all Fs < 1.6, ps $> .2$). Mean rate of recall across groups was $M = 0.60$ ($SE = 0.02$). A parallel analysis of the retraction item among the fact questions yielded only the trivial main effect of retraction, $F(2, 90) = 4.85$, $p = .01$. Rates of retraction recall were higher, $M = 0.58$ ($SE = 0.06$) if there was a retraction than if there was none, $M = 0.25$ ($SE = 0.08$). We conclude that the emotionality groups did not differ in their overall memory abilities and that the scenarios did not differ in general memorability.

References to initial misinformation

References to initial misinformation were corrected using the alternative-throughout control group's baseline (i.e., the mean number of references spontaneously made to bad weather or a terrorist attack without this information having been presented), which unlike Experiment 1 was different from zero (0.22 and 0.06 for bad weather and terrorist attack, respectively).

The mean corrected misinformation scores are shown in the top panel of Table 3. The results attest to the success of our retention interval manipulation because the overall levels of reference to initial misinformation were considerably lower than those in Experiment 1.

A 2×3 ANOVA on baseline-corrected references to misinformation revealed a main effect of retraction, $F(2, 90) = 12.91$, $\eta^2 = .22$,

Table 3. *Mean number of references to initial misinformation in Experiment 2*

		Emotionality					
		Low		*High*		*Combined*	
Sample	*Retraction condition*	*M*	*SE*	*M*	*SE*	*M*	*SE*
All participants	No-retraction	2.75	0.38	2.72	0.70	2.73	0.39
	Retraction-only	1.16	0.50	0.56	0.30	0.86	0.29
	Causal-alternative	0.78	0.26	0.75	0.32	0.77	0.20
Subsample recalling retraction	Retraction-only	0.73[a]	0.54[a]	0.60[a]	0.33[a]	0.66	0.31
	Causal-alternative	0.56[a]	0.22[a]	0.84[b]	0.45[b]	0.71	0.26

Note: Number of participants in each cell is $N = 16$ unless indicated otherwise by superscripts.
Maximum score is 24.
[a]$N = 9$. [b]$N = 10$.

Table 4. *Mean number of references to correct information in Experiment 2*

	Emotionality					
	Low		*High*		*Combined*	
Retraction condition	*M*	*SE*	*M*	*SE*	*M*	*SE*
No-retraction	0.13	0.09	0.16	0.11	0.14	0.07
Retraction-only	0.03	0.03	0.06	0.06	0.05	0.03
Causal-alternative	3.00	0.44	3.41	0.62	3.20	0.38

Note: Number of participants in each cell is $N = 16$, and maximum score is 24.

$p < .001$, but no other effects (Fs < 1). Follow-up HSD tests showed that the no-retraction conditions differed from both the retraction-only and the causal-alternative conditions at both levels of emotionality (ps $< .001$; there was no reliable difference between the two latter conditions). The rate of references to initial misinformation was again larger than zero even in the causal-alternative condition, $t(31) = 3.82$, $p = .001$. Planned comparisons showed that were there no reliable differences between levels of emotionality at any level of the retraction variable (all Fs < 1).

Restricting this analysis to those participants who explicitly remembered the retraction ($N =$ 37 out of 64 participants in the four retraction conditions) yielded exactly the same pattern (see bottom panel of Table 3; interaction of emotionality and retraction, $F < 1$).

References to the alternative information

References to the alternative information (i.e., faulty fuel tank) are presented in Table 4 (the alternative-throughout control condition, $M =$ 2.81, $SE = 0.41$, is not shown in the table). As in Experiment 1, this information was only given to the causal-alternative conditions and the alternative-throughout control condition.

Accordingly, there were few references to the (nonpresented) alternative information in the no-retraction and retraction-only conditions, and those data were not considered further. The causal-alternative conditions were entered into a one-way ANOVA together with the control condition. The analysis revealed no significant effect, $F < 1$. Planned comparisons showed that emotionality had no effect in the causal-alternative condition, $F < 1$.

Table 5. *Effect sizes from selected studies reporting an influence of emotionality on memory*

Study	*Exp.*	*Effect*	d
Kensinger & Corkin (2003a)	1	Word recollection negative > neutral	0.92
	2	Word recognition negative > neutral	0.63
	2	Word source memory negative > neutral	0.81
	3	Word recollection negative > neutral	0.64
	4	Word recognition negative > neutral	1.25
	4	Word source memory negative > neutral	0.7
	5	Word recollection negative > neutral	0.54
	6	Word recall negative > neutral	0.66
Kensinger (2008)	2	Word recognition negative > neutral	0.72
Heuer & Reisberg (1990)		Central story information recall negative > neutral	1.07
		Slide recognition negative > neutral	1.51
		Sentence recognition negative > neutral	1.21
Hamann, Ely, Grafton, & Kilts (1999)		Picture recognition negative > neutral	2.15

Note: d = effect size. Exp. = experiment.

Power analyses

To appraise the statistical power of the present experiment to actually detect a potential effect of emotion, we carried out power analyses (using G*power 3; Faul, Erdfelder, Lang, & Buchner, 2007; following suggestions by Dunlap, Cortina, Vaslow, & Burke, 1996). Specifically, we asked how likely it would have been for the present experiment to detect an effect of the size observed in other relevant research.

To this end, we first surveyed the effect sizes in published studies that reported effects of the emotionality of the memoranda on performance (see Table 5 for a representative summary of studies that included sufficient information to permit computation of effect sizes). We focused on the study by Kensinger and Corkin (2003a) because it, like ours, used verbal materials and provided multiple measurements. Compared to other studies, effect sizes in this experiment were relatively moderate, with the average $d = 0.77$. This parameter can be transformed into the ANOVA-associated effect-size parameter f (cf. Faul et al., 2007), yielding $f = 0.39$. The power of the present experiment to detect an effect of that magnitude is .93 (where the effect is the interaction between emotionality and retraction).

We next conducted a within-experiment comparison by considering our manipulation check of the emotionality of the material. The manipulation check yielded significant effects of similar size, with $d = 0.84$ ($f = 0.43$; VAS sadness) and $d = 0.72$ ($f = 0.34$; PANAS), respectively. The size of these effects confirms that our emotionality manipulation was strong enough to elicit effects that were comparable in magnitude to relevant precedents. Moreover, had these effects transferred from the self-report of emotion to people's memory performance with the same materials within the same experiment, the power of Experiment 2 to detect an Emotionality × Retraction interaction would have been .97 (based on VAS sadness) and .84 (PANAS), respectively.

We conclude that this power analysis provides strong evidence that Experiment 2 was sufficiently powerful to detect emotionality effects of the magnitude typically reported in studies in which emotionality of verbal material influenced memory. Moreover, the experiment was sufficiently powerful to detect an effect of emotionality of the same magnitude that was observed with the exact same stimuli in the same experiment during self-report of their emotional impact.

Discussion

Experiment 2 successfully replicated the main result of Experiment 1: We again found no suggestion that emotionality determines the continued influence of misinformation, despite increased statistical power, a reduction in references to initial misinformation following retraction alone (thus removing a possible ceiling effect), and demonstrable emotional impact of the scenarios. The latter finding is particularly relevant because it confirms that memory updating took place at a time when participants were demonstrably emotional. In particular, the manipulation check showed that high- and low-emotionality scenarios differed not only in content emotionality (i.e., deliberate killing is undoubtedly more emotional than an accident; cf. Unz et al., 2008) but also in their impact on participants' self-reported feelings. Notwithstanding, emotionality did not influence the continued influence of misinformation.

In contrast to Experiment1, however, in Experiment 2 the retraction of initial misinformation alone—even if not accompanied by a causal alternative—was sufficient to reduce reliance on initial misinformation. Although this is in contrast with Experiment 1, there are also precedents showing that in some cases a mere retraction is sufficient to reduce—but not eliminate—the continued influence effect (Bush, Johnson, & Seifert, 1994, reviewed in Seifert, 2002; Ecker et al., in press; Wilkes & Reynolds, 1999). The reasons for these inconsistent findings are currently unclear (and outside of the focus of the present article). Most important for our purposes, retention interval did not interact with our manipulation of greatest interest—namely, emotionality.

Our finding that emotionality did not influence the recall of peripheral detail is not very surprising. As discussed in the introduction, the positive effects of emotion on memory are largely confined to central aspects of an event (cf. Heuer & Reisberg, 1990).

One remaining concern may be that the two plane crash scenarios compared in Experiments 1 and 2 both involved many fatalities and that therefore both may have been perceived as emotionally negative events, thus reducing the efficacy of our main manipulation. The final experiment therefore added a low-emotionality scenario with no fatalities to allow comparison with a more neutral baseline.

Moreover, because the factor of greatest interest—emotionality—gave rise to a null effect in Experiments 1 and 2, we (a) again increased sample size and hence power and (b) sought to enhance the impact of the emotionality manipulation by leading participants to believe that the report referred to a real-world incident.

EXPERIMENT 3

Experiment 3 was similar to Experiment 2 but used three levels of emotionality and hence a 3 × 3 between-subjects design augmented by the usual alternative-throughout control group.

Method

Participants

A total of 200 second-year psychology students participated in this experiment for partial course credit (62 males, mean age 20.0 years, age range 18–37 years). None of them had taken part in Experiment 1 or 2. An equal number of participants ($N = 20$) were randomly assigned to each of the 10 conditions.

Stimuli

The scenarios again reported a plane crash. In the low- and medium-emotionality conditions, the plane was initially reported to have crashed due to bad weather. The high-emotionality scenario differed from the medium-emotionality scenario in that the cause of the plane crash was again a terrorist attack (on-board shooting). The conditions differed in the number of fatalities, which was 0 in the low-emotionality condition and in excess of 100 in the medium- and high-emotionality conditions. Also, we again chose departure cities that are or are not frequently associated with terrorism: In the low- and medium-emotionality scenarios, the fictitious plane departed in

Athens, Greece. In the high-emotionality condition, the plane departed in Tehran, Iran, to enhance the terrorism innuendo (cf. Footnote 3). In all conditions, the plane was headed for London, crashing near Strasbourg in north-eastern France.

Some modifications were made to the materials to make sure that the scenarios were all plausible but as similar as possible. For example, the plane now exploded after an attempted emergency landing, to provide a plausible scenario for the survival of all passengers in the low-emotionality condition; this also implies that the reported explosion in all scenarios followed from the impact of the emergency landing, not directly from the actual crash cause (i.e., on-board shooting, lightning strike, fuel system malfunction). Scenarios hence differed mainly in what caused the plane to stall and made the emergency landing necessary. Participants were told that the reports were edited newspaper excerpts from several 1989 issues of the Sydney Morning Herald, thus creating the impression that the scenario reported an actual plane crash. (It is for this reason that the "crash site" was moved to Europe as participants would immediately identify any major crash near Perth to be fictitious.) After the experiment, participants were fully debriefed about the fictitious nature of the scenarios.

To verify the impact of the emotionality manipulation, all participants received the happy/sad and angry/calm VAS and the PANAS immediately after reading the reports. In the VAS, participants who read the highly emotive scenario rated themselves as sadder ($M = 7.0$ cm; $SE = 0.2$) and angrier ($M = 4.9$ cm; $SE = 0.2$) than people who read the medium-emotionality (sadness: $M = 6.5$ cm; $SE = 0.2$; anger: $M = 4.5$ cm; $SE = 0.2$) or low-emotionality scenarios (sadness: $M = 5.5$ cm; $SE = 0.2$; anger: $M = 3.8$ cm; $SE = 0.2$. One-way ANOVAs showed the effect to be highly significant for both sadness, $F(2, 177) = 16.24$, $\eta^2 = .155$, $p < .001$, and anger, $F(2, 177) = 5.54$, $\eta^2 = .059$, $p < .01$ (the effect between medium- and high-emotionality was marginally significant for sadness, $F(1, 118) = 3.29$, $\eta^2 = .027$, $p = .07$).

On the PANAS, participants who read the high-emotionality scenario scored higher on the negative affect scale ($M = 18.62$, $SE = 1.02$) than people who read the medium- ($M = 16.03$; $SE = 0.59$) or low-emotionality scenarios ($M = 13.22$; $SE = 0.45$). This was a highly significant effect, $F(2, 177) = 13.76$, $\eta^2 = .135$, $p < .001$; the difference between medium and high emotionality by itself was also significant, $F(1, 118) = 4.81$, $\eta^2 = .039$, $p < .05$. As in Experiment 2, there was no difference on the positive affect scale (emotionality high: $M = 16.03$, $SE = 0.51$; medium: $M = 15.47$, $SE = 0.44$; low: $M = 15.32$; $SE = 0.58$), $F < 1$. A further ANOVA found the differences of the difference scores—positive minus negative affect—also to be highly significant, $F(2, 177) = 8.89$, $\eta^2 = .091$, $p < .001$.

Hence, we again conclude that our scenarios successfully instantiated the desired emotionality manipulation.

Procedure

The procedure was identical to Experiment 2 with the following exceptions. First, participants were tested in groups during a tutorial session (average group size was 11.3 participants, ranging from 2 to 18). The retention interval was now approximately 20 minutes, during which participants attended to an unrelated presentation.

Results

Coding of responses

Responses were again scored by a trained scorer who was blind to condition. A further two independent judges scored 30 (3 per condition) randomly selected scoring sheets each. This audit confirmed that the inter-rater reliability for causal references was high ($r = .91$).

Fact recall

A 3 × 3 ANOVA on fact recall revealed no significant effects (all Fs < 2.2, ps > .1). Mean rate of recall across groups was $M = 0.45$ ($SE = 0.02$). Close inspection of all conditions suggested, however, that fact recall was selectively impaired in

the high-emotionality condition featuring a causal alternative ($M = 0.33$, $SE = 0.05$; in other conditions M ranged from 0.43 to 0.51). The effect was significant in a contrast analysis, contrasting this condition with all others, $F(1, 171) = 5.22$, $p = .02$.

A parallel 3 × 3 analysis of the retraction item among the fact questions yielded only the trivial main effect of retraction, $F(2, 171) = 32.86$, $p < .001$. Rates of retraction recall were higher, $M = 0.73$ ($SE = 0.04$) if there was a retraction than if there was none, $M = 0.18$ ($SE = 0.05$). We conclude that the emotionality groups did not generally differ in their overall memory abilities—although the high-emotionality condition with a causal alternative suggested that memory for peripheral details was negatively affected by emotionality.

References to initial misinformation

References to initial misinformation were corrected using the alternative-throughout control group's baseline (i.e., the mean number of references spontaneously made to bad weather or a terrorist attack without this information having been presented), which was 0.10 for bad weather and 0.15 for a terrorist attack.

The mean corrected misinformation scores are shown in the top panel of Table 6.

A 3 × 3 ANOVA on baseline-corrected references to misinformation revealed a main effect of retraction, $F(2, 171) = 57.44$, $\eta^2 = .40$, $p < .001$, but no other effects (Fs < 1.26). Follow-up HSD tests showed that the no-retraction conditions differed from both the retraction-only and the causal-alternative conditions at all three levels of emotionality (ps < .02). Retraction-only and causal-alternative conditions (averaging across emotionality) also differed significantly as indicated by a planned contrast, $F(1, 114) = 7.12$, $\eta^2 = .059$, $p < .01$. Again, the rate of references to initial misinformation was larger than zero even in the causal-alternative condition, $t(179) = 10.13$, $p < .001$. Importantly, there were no significant differences between levels of emotionality at any level of the retraction variable (all Fs < 1.53, all ps > .2).

Restricting the omnibus analysis to those participants who explicitly remembered the retraction ($N = 87$ out of 120 participants in the six retraction conditions) yielded exactly the same pattern (see bottom panel of Table 6; interaction of emotionality and retraction, $F < 1$).

References to the alternative information

References to the alternative information (i.e., faulty fuel tank) are presented in Table 7 (the alternative-throughout control condition, $M = 4.15$, $SE = 0.31$, is not shown in the table). As in Experiments 1 and 2, this information was only given to the causal-alternative conditions and the alternative-throughout control condition. These were entered into a one-way ANOVA, which showed a main effect of condition, $F(3, 76)$

Table 6. *Mean number of references to initial misinformation in Experiment 3*

		Emotionality							
		Low		*Medium*		*High*		*Combined*	
Sample	*Retraction condition*	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
All participants	No-retraction	3.60	0.42	2.63	0.49	3.48	0.53	3.23	0.28
	Retraction-only	0.90	0.33	0.95	0.26	0.90	0.34	0.92	0.18
	Causal-alternative	0.28	0.18	0.62	0.23	0.18	0.14	0.36	0.11
Subsample recalling retraction	Retraction-only	0.86[a]	0.49[a]	1.07[a]	0.35[a]	0.66[b]	0.33[b]	0.84	0.22
	Causal-alternative	0.33[c]	0.24[c]	0.45[d]	0.20[d]	0.18[c]	0.16[c]		

Note: Number of participants in each cell is $N = 20$ unless indicated otherwise by superscripts. Maximum score is 24.
[a]$N = 12$. [b]$N = 16$. [c]$N = 15$. [d]$N = 17$.

Table 7. *Mean number of references to correct information in Experiment 3*

	Emotionality							
	Low		*Medium*		*High*		*Combined*	
Retraction condition	*M*	*SE*	*M*	*SE*	*M*	*SE*	*M*	*SE*
No-retraction	0.05	0.05	0.05	0.05	0.08	0.05	0.06	0.03
Retraction-only	0.20	0.16	0.40	0.24	0.05	0.05	0.22	0.10
Causal-alternative	3.33	0.37	4.28	0.56	2.43	0.40	3.34	0.28

Note: Number of participants in each cell is $N = 20$, and maximum score is 24.

$= 4.11$, $\eta^2 = .140$, $p = .01$, with fewer references to the alternative information in the high- than in the medium-emotionality and alternative-throughout control conditions, Fs(1, 38) > 7.15, $\eta^2 > .158$, $p < .011$.

Power analyses

The power of Experiment 3 to detect an effect of the magnitude found in Kensinger and Corkin (2003a; $d = 0.77$, $f = 0.39$) is .99 (where the effect is the interaction between emotionality and retraction).

We again examined the effect sizes of our manipulation check regarding the emotionality of the material, which were $d = 0.98$ ($f = 0.49$; VAS sadness) and $d = 0.89$ ($f = 0.44$; PANAS negative affect scale), respectively. We note that the size of these effects confirms that our emotionality manipulation was strong enough to elicit substantial effects, comparable in magnitude to relevant precedents. Had these effects transferred to the same people's comprehension performance with the same materials within the same experiment, the power of Experiment 3 to detect an Emotionality × Retraction interaction would have been .99 (based on either VAS or PANAS).

We suggest that this power analysis provides strong evidence that Experiment 3 was sufficiently powerful to detect emotionality effects of the magnitude typically reported in relevant precedents, or that were observed with the exact same stimuli in the same experiment during self-report of their emotional impact.

Discussion

Experiment 3 successfully replicated the main result of Experiments 1 and 2: We again found no suggestion that emotionality determines the continued influence of misinformation, despite increased statistical power, addition of a more neutral low-emotionality baseline condition, and substantial emotional impact of the scenarios. The manipulation check demonstrated that high- and low-emotionality scenarios differed considerably in their impact on participants' self-reported feelings. Nevertheless, emotionality clearly did not influence the continued influence of misinformation.

The fact that one condition—the high-emotionality condition with a causal alternative—exhibited significantly worse recall of corrective information and peripheral facts provides a hint that memory for details may have been impaired by high emotionality. Because these effects were only apparent in a single cell, caution in interpretation is certainly advised. Yet, a negative effect of emotionality on memory for details would be entirely in line with previous research (e.g., Burke et al., 1992; Schmidt, 2004).

GENERAL DISCUSSION

We asked a very simple but important question: Does the emotionality of material determine people's ability to discount initial misinformation when it is later retracted? Three studies provided a clear and unambiguous answer: Notwithstanding

the demonstrable success of our emotionality manipulation to induce negative affect, this manipulation did not alter the way in which people processed initial misinformation. To underscore the consistency of these results, Figure 1 summarizes the data of all three experiments; it shows the effects of emotionality—comparing the lowest and highest level of emotionality used in each experiment—on the continued influence of misinformation across the three levels of retraction. An omnibus analysis of these data across experiments yields sufficient power to detect an effect of the magnitude $f = 0.39$ or larger (Kensinger & Corkin, 2003a) with a probability of .9998 and a power to detect even a rather moderate effect (e.g., $f = 0.25$) with a probability of .97.

Potential limitations and objections

Although this power analysis should be more than sufficient to allay fears about acceptance of the null hypothesis, we sought further support for our conclusion through a Bayesian analysis of our data. Conventional frequentist test statistics, such as those used in the preceding Results sections, have been the standard mode of analysis over the past decades, but the null-hypothesis-testing approach has not been without its critics (Gallistel, 2009; Killeen, 2005; Krueger, 2001; Tryon, 2001; Wagenmakers, 2007). One of the serious drawbacks of a frequentist approach is its focus on rejection of the null hypothesis and its inability to provide much guidance for the assessment of null results (for a recent critique, see Wagenmakers, Lee, Lodewyckx, & Iverson, 2008). In contrast, recently developed Bayesian techniques allow one to estimate and rationally quantify the statistical evidence in favour of the null hypothesis (e.g., Rouder, Speckman, Sun, & Morey, 2009; Wagenmakers, 2007; Wetzels, Raaijmakers, Jakab, & Wagenmakers, 2009). Under a Bayesian approach, the interpretative asymmetry that besets null hypothesis testing can be avoided, and strong conclusions about null effects become possible and permissible. We therefore used a Bayesian approach to hypothesis testing to calculate the posterior probability, given our data from each experiment, that there was no interaction between emotionality and the retraction variable. This test was based on the Bayesian information criterion (BIC) for statistical inference (cf. Raftery, 1995; Wagenmakers, 2007).

In a nutshell, BIC is a well-known statistic that is widely used for model comparison in experimental psychology (e.g., Heathcote, Raymond, & Dunn, 2006; Oberauer & Bialkova, 2009). In the context of data analysis, BIC is computed for two models: one that contains the effect of interest (e.g., a main effect or interaction in an ANOVA design) and one that does not. The difference between those BICs can then be used to derive a Bayes factor (BF) approximation (Wagenmakers, 2007). The BF is defined as the ratio of prior predictive probabilities—that is, the ratio of the probability of observing the data D given the null hypothesis H_0 over the probability of the data given H_1, or $\Pr(D|H_0)/\Pr(D|H_1)$. For example, a BF of 20 implies that the observed data are around 20 times more likely to occur under H_0 than under H_1. When H_0 and H_1 are assumed to be equally likely a priori, which is a standard assumption that reflects complete uncertainty, such a result would imply that the posterior probability in favour of the null hypothesis given the data, $\Pr(H_0|D)$, is $20/21 = .95$. According to the classic guidelines of Jeffreys (1939/1998), a BF between 10 and 30 would be considered "strong" evidence in favour of the null hypothesis, a BF between 30 and 100 would constitute "very strong" evidence, and a $BF > 100$—with $\Pr(H_0|D) > .99$—would be interpreted as "decisive" evidence. Conversely, a BF below 1 constitutes evidence in favour of the alternative hypothesis along a similarly graded scale (i.e., the smaller the BF, the stronger the evidence for the alternative hypothesis). Thus, in contrast to null-hypothesis testing, BICs are entirely symmetrical and permit a quantification of the weight of evidence in favour of the null hypothesis as well as against it.

We calculated BICs and BFs in this manner for all effects in each of the experiments. Our analyses used the BIC function of the nlme package (Pinheiro, Bates, DebRoy, Sarkar, & the R Core Team, 2009) of the R programming language

Table 8. Bayes factors and posterior probabilities for Experiments 1–3

Experiment	*Effect*	*BF*	$Pr(H_0\|D)$	$Pr(H_1\|D)$
1	Retraction	<0.001		>.999
	Emotionality	0.93	.48	
	Interaction	37.23	.97	
2	Retraction	<0.001		>.999
	Emotionality	8.02	.89	
	Interaction	71.58	.99	
3	Retraction	<0.001		>.999
	Emotionality	140.37	>.99	
	Interaction	2,358.08	>.999	
1–3 combined	Retraction	<0.001		>.999
	Emotionality	16.61	.94	
	Interaction	263.91	>.99	

Note: Bayes factors and posterior probabilities of H_0 and H_1 given the data [$Pr(H_0|D)$ and $Pr(H_0|D)$], where Pr = probability, D = data, H = hypothesis. BF = Bayes factor.

(R Development Core Team, 2008); some statistics were derived from the ANOVA table as demonstrated by Wagenmakers (2007, p.798–799). We report the BFs together with the posterior probabilities for all three experiments in Table 8. The table shows the Bayesian statistics for the effect of prime interest—the Retraction × Emotionality interaction—and the two constituent main effects. For each effect, we show the posterior probability of the more likely hypothesis (as indicated by the ANOVA results), which is H_1 for the main effect of retraction, but H_0 for both the main effect of emotionality and the two-way interaction (this follows precedent, see Figure 8 of Gallistel, 2009).

Table 8 shows that Experiments 1 and 2 each provided very strong evidence in favour of a purely additive model with no interaction of emotionality and retraction; Experiment 3 provided "decisive" evidence. The table also shows that our Bayesian analysis confirmed the ANOVA results with respect to the main effect of retraction in all three experiments.

The last panel of Table 8 reports the results of a conjoint analysis of all three experiments, comparing the lowest and highest levels of emotionality used in each experiment (as shown in Figure 1). Across experiments, this analysis yielded decisive evidence for a main effect of retraction and against an interaction of emotionality and retraction and also strong evidence against a main effect of emotionality.

We conclude that the null effect of the crucial Retraction × Emotionality interaction observed in three studies, with a total of more than 380 participants, is highly unlikely to reflect a lack of statistical power. Instead, the results provide demonstrably and quantifiably strong evidence in favour of the null hypothesis.

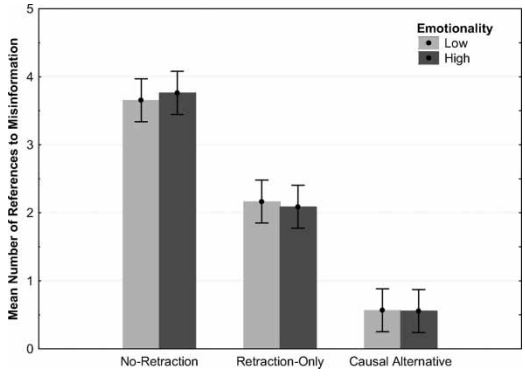

Figure 1. *The effect of emotionality on the mean number of references to initial misinformation across levels of retraction. Data are collapsed across Experiments 1–3 (comparing the lowest and highest levels of emotionality used in each experiment); hence each data point is based on 46 participants. Error bars show standard errors of the mean.*

Another objection that might be raised about our studies concerns the strength of our manipulation of emotionality. Perhaps our terrorism scenario was insufficiently emotive? Alternatively, any plane crash is a disastrous event irrespective of its causes, so perhaps even our low-emotionality condition was highly emotive, thus negating any additional selective effect of the terrorism scenario? We cite three arguments against those potential objections.

First, Experiment 3 included a scenario in which the plane crash led to no fatalities, thus presumably creating a more "neutral" scenario than the low-emotionality scripts used in Experiments 1 and 2. (In support, the subjective ratings of affect for the neutral scenario in Experiment 3 were below those for the low-emotionality scenario in Experiment 2, which included a large number of fatalities.) Nonetheless, the neutral scenario in Experiment 3 gave rise to a pattern of discounting of initial misinformation that was identical to that observed with the maximally emotive terrorism scenarios.

Second, we directly confirmed in Experiments 2 and 3 that our scenarios were sufficiently different and powerful to induce demonstrably different levels of negative affect. This puts to rest any concern about a potential lack of efficacy of our manipulation. The simultaneous presence of an effect of emotionality—on ratings of affect—and its absence—on retraction—further supports the notion that the latter null result cannot be ascribed simply to a lack of power.

Third, we stress that our manipulation was chosen to resemble the likely impact of news stories in the daily media on people's affect and behaviour (cf. Gross & D'Ambrosio, 2004; Jonathan, Alison, & Long, 2007; Sinaceur, Heath, & Cole, 2005). In particular, Healy et al. (2009) found that even obviously fictitious news reports of repeated intentional violence can have escalating effects on people's feelings. Unz et al. (2008) investigated the effects of real TV news on self-reported affect and found significant differences between news that contained intentional violence against humans (e.g., terrorist attacks) and reports of unintentional violence against humans (e.g., accidents). For this comparison, they reported effect sizes (e.g., $d = 0.99$ for anger and $d = 0.57$ for sadness) that were absolutely comparable in size to the effects we found in the present study. We conclude that our manipulations had emotional effects equivalent to the effects of real news. We stress again that we were not interested in the effects of high-level emotion as, for instance, those following life-threatening trauma. However, this should not imply that moderate emotionality cannot have significant consequences. For example, Veno and van den Eynde (2007) demonstrated how media reports can lead to long-term changes in public opinion by fuelling or mitigating emotions.

We conclude that our manipulation provided a firm basis for our conclusion that moderate negative emotionality of material similar to real news reports did not affect people's discounting of initial misinformation. Having dealt with possible objections to our results, we now discuss their theoretical implications.

The effectiveness of a mere retraction

In the first experiment, people continued to rely on initial misinformation even after it was retracted unless the retraction was accompanied by a causal alternative (replicating H. M. Johnson & Seifert, 1994). In the second and third experiments, a retraction proved sufficient to reduce reliance on the initial misinformation (although provision of causal alternative further reduced that reliance at least in Experiment 3; this replicates Bush et al., 1994, and Wilkes & Reynolds, 1999). The most obvious difference between our experiments was the retention interval, which was short in Experiment 1 (5 min) and longer in Experiments 2 and 3 (40 and 20 min, respectively). The other studies mentioned all had rather short retention intervals, although Bush et al. and Wilkes and Leatherbarrow (1988) did not specify these specifically in their reports. In order to test whether differences in retention interval may mediate the effects of a mere retraction, we conducted another experiment using the low-emotionality materials of Experiment 3, with a 1-minute retention interval. Even with

this minimal retention interval, we found that a mere retraction significantly reduced the continued influence effect. Hence it is currently unclear why a mere retraction is sometimes sufficient to reduce the continued influence of misinformation and sometimes not. We speculate that one relevant factor could be subtle differences between the scenarios in the ability of the postretraction account to explain specific features of the event. That is, the success of a retraction may hinge on the severity of the event-model gap that the retraction leaves behind, and people should be more likely to rely on misinformation the more specific event features are left unexplained by the retraction (cf. also Ecker et al., in press). However, given the similarity of scenarios across experiments in the present study, it is unclear whether this could be a viable explanation, and these speculations remain to be addressed by future research. Most importantly in the present context, retention interval did not interact with emotionality. Also, observing equivalent effects of emotionality across various retention intervals refutes the notion that emotion may only become effective during the stage of consolidation of the scenario's corrected representation (cf. Burke et al., 1992; McGaugh, 2000; Sharot & Yonelinas, 2008).

Theoretical implications

As discussed in the introduction, the dominant view in the current literature is that negative emotionality leads to more stable memory representations of central stimulus aspects (cf. Kensinger, 2009; Levine & Pizarro, 2004). It is reasonable to assume that more stable representations should be more resistant to modification. Therefore, the present results may appear quite surprising. However, as mentioned at the outset, while it is clear that emotion can enhance memory, claims that emotion generally enhances memory have been challenged. For instance, Cubelli and Della Sala (2008) have argued that in many cases it is not emotion itself but the salience of the material and the consequent absence of interference by similar events that causes the strong recollections associated with emotive material. Accordingly, Dougal and Rotello (2007) found that while emotion indeed increased subjective reports of "remembering" (i.e., conscious recollection) in recognition memory tasks, it did not enhance recollection or sensitivity per se. Instead, some reported that effects of emotionality seem to arise from changes in response bias (see also Langeslag & van Strien, 2008; Windmann & Kutas, 2001). Talarico and Rubin (2003) similarly demonstrated that emotional memories are characterized by elevated levels of confidence (associated with vividness and belief in their accuracy) rather than by improved accuracy and consistency. Hence, it seems that emotion enhances accuracy primarily for the mere fact that an event occurred, but beyond that, it mainly boosts the recollective experience rather than its accuracy (Phelps & Sharot, 2008). Interestingly, this recollective experience associated with emotional material does not show the usual decline over time, whereas the decline in actual recollection accuracy over time is equivalent for emotional and neutral materials (Sharot & Yonelinas, 2008). The latter point was nicely demonstrated by a study showing that memories of 9/11 deteriorated at the same rate as less emotional memories (Talarico & Rubin, 2007). Our results—including the no-retraction conditions, which provide a baseline measure of how well people processed the material—show no evidence of memory strengthening elicited by the emotionality of the material, and our results are thus in line with precedents.

Furthermore, we found no evidence that emotionality determines memory updating. While ours is the first study to investigate this issue in a correction-of-misinformation paradigm, it is noteworthy that Kensinger and Corkin (2003b) likewise found no influence of emotionality on updating of working memory. Participants in their study performed a variety of working-memory tasks with emotional and neutral stimuli. One of the tasks was a self-ordered pointing task—participants were presented with a grid of 15 pictures 15 times in random arrangements and were asked to point to a different picture every time. This task requires continuous updating of the set of already selected pictures. Performance did not depend on whether

the pictures were emotive or neutral. Likewise, emotionality had no robust or consistent effect on performance on an n-back task (using emotional versus neutral faces and words), which requires the continuous updating of the activated information in working memory.

Taken together, our results imply that there are domains of mnemonic processing that are unaffected by emotionality. Despite the convincing demonstrations of the effects that emotions can have on memory—both positive (e.g., Conway et al., 1994; D'Argembeau & van der Linden, 2004; the recall and source memory experiments in Kensinger & Corkin, 2003a) and negative (e.g., Burke et al., 1992; Christianson & Loftus, 1987)—these effects are not ubiquitous. Moreover, the correction of initial misinformation as investigated in the present study is not the only memory-related process that is apparently unaffected by emotion. Evidence is beginning to accrue that other, related memory functions such as directed forgetting are likewise not influenced by emotional content valence (Barnier et al., 2007; McNally, Clancy, Barrett, & Parker, 2004). Both these memory functions involve the intentional retraction of information, and this mechanism is apparently not affected by emotionality. Theories of memory and emotion that propose tight links between emotion and memory (Kensinger, 2007; Levine & Pizarro, 2004) need to take this into account.

Implications outside the laboratory

Our findings have implications for "real-world" issues. We presented information to participants in a way that resembles the typical presentation of information in newspapers and the media in general. Indeed, real news reports contain corrections and retractions in themselves very frequently. For example, a New York Times article published on January 31, 2007—titled "Iran May Have Trained Attackers That Killed 5 American Soldiers, U.S. and Iraqis Say"—first states that "attackers who used American-style uniforms and weapons to infiltrate a secure compound and kill five American soldiers in Karbala on Jan. 20 may have been trained and financed by Iranian agents, according to American and Iraqi officials knowledgeable about the inquiry" (Glanz & Mazzetti, 2007, para 1), only to state in the following paragraph that "Officials cautioned that no firm conclusions had been drawn and did not reveal any direct evidence of a connection" (Glanz & Mazzetti, 2007, para 2). Similarly, on March 22, 2003, CNN reported "UK: Thousands of Iraqis surrender", stating that according to a British official, "the 51st Division of the Iraqi army—with 8,000 to 10,000 soldiers—surrendered in the Basra area" (CNN, 2003, para 2). This was qualified two paragraphs later by citing a U.S. official that "between only one and two thousand Iraqi troops are in custody" (CNN, 2003, para 4). Our findings suggest that the processes involved in such updating of information are equally effective (or equally error prone, depending on retention interval or on whether or not a causal alternative is provided) irrespective of the emotionality of the information.

In closing, it is important to note that no study has yet convincingly demonstrated a complete annulment of the influence of initial misinformation, even if the corrections were clear, memorable, immediately followed the initial misinformation, and supplied causal alternatives (e.g., Ecker et al., in press; H. M. Johnson & Seifert, 1994; van Oostendorp & Bonebakker, 1999). In the present study, significant levels of false inferences persisted in all conditions, even when a correction was successful (i.e., it reduced the influence of initial misinformation compared to no retraction). Newspaper articles such as those mentioned above—which qualify the message initially delivered in the headline—are therefore virtually designed to produce false inferences, especially immediately upon reading. Thus, although the continued influence of misinformation may not be stronger with emotive than with more neutral information, it constitutes a powerful effect that is readily elicited by many media reports, whether inadvertently or not.

Original manuscript received 01 September 2009
Accepted revision received 25 February 2010
First published online 06 August 2010

REFERENCES

Barnier, A. J., Conway, M. A., Mayoh, L., Speyer, J., Avizmil, O., & Harris, C. B. (2007). Directed forgetting of recently recalled autobiographical memories. *Journal of Experimental Psychology: General*, *136,* 301–322.

Brainerd, C. J., Stein, L. M., Silveira, R. A., Rohenkohl, G., & Reyna, V. F. (2008). How does negative emotion cause false memories? *Psychological Science*, *19,* 919–925.

Brown, R., & Kulik, J. (1977). Flashbulb memories. *Cognition*, *5,* 73–99.

Bullinaria, J. A., & Levy, J. P. (2007). Extracting semantic representations from word co-occurrence statistics: A computational study. *Behavior Research Methods*, *39,* 510–526.

Burke, A., Heuer, F., & Reisberg, D. (1992). Remembering emotional events. *Memory & Cognition*, *20,* 277–290.

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: Lawrence Erlbaum Associates.

Cainkar, L. (2002). No longer invisible: Arab and Muslim exclusion after September 11. *Middle East Report*, *224,* 22–29.

Christianson, S.-Å. (1989). Flashbulb memories: Special, but not so special. *Memory & Cognition*, *17,* 435–443.

Christianson, S.-Å., & Loftus, E. F. (1987). Memory for traumatic events. *Applied Cognitive Psychology*, *1,* 225–239.

Christianson, S.-Å., & Loftus, E. F. (1990). Some characteristics of people's traumatic memories. *Bulletin of the Psychonomic Society*, *28,* 195–198.

Christianson, S.-Å., & Loftus, E. F. (1991). Remembering emotional events: The fate of detailed information. *Cognition & Emotion*, *5,* 81–108.

CNN. (2003, March 22). *UK: Thousands of Iraqis surrender*. Retrieved October 21, 2008, from http://edition.cnn.com/2003/WORLD/europe/03/22/sprj.irq.britain.mod/index.html

Conway, M. A., Anderson, S. J., Larsen, S. F., Donnelly, C. M., McDaniel, M. A., McClelland, A. G. R., et al. (1994). The formation of flashbulb memories. *Memory & Cognition*, *22,* 326–343.

Crombag, H. F. M., Wagenaar, W. A., & van Koppen, P. J. (1996). Crashing memories and the problem of "source monitoring". *Applied Cognitive Psychology*, *10,* 95–104.

Cubelli, R., & Della Sala, S. (2008). Flashbulb memories: Special but not iconic. *Cortex*, *44,* 908–909.

D'Argembeau, A., & van der Linden, M. (2004). Influence of affective meaning on memory for contextual information. *Emotion*, *4,* 173–188.

de Vega, M., Urrutia, M., & Riffo, B. (2007). Canceling updating in the comprehension of counterfactuals embedded in narratives. *Memory & Cognition*, *35,* 1410–1421.

Doerksen, S., & Shimamura, A. P. (2001). Source memory enhancement for emotional words. *Emotion*, *1,* 5–11.

Dougal, S., & Rotello, C. M. (2007). "Remembering" emotional words is based on response bias, not recollection. *Psychonomic Bulletin & Review*, *14,* 423–429.

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

Ecker, U. K. H., Lewandowsky, S., & Tang, D. T. W. (in press). Explicit warnings reduce but do not eliminate the continued influence of misinformation. *Memory & Cognition*.

Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition & Emotion*, *6,* 409–434.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39,* 175–191.

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 and Social Psychology Bulletin*, *23,* 1215–1226.

Gallistel, C. R. (2009). The importance of proving the null. *Psychological Review*, *116,* 439–453.

Glanz, J., & Mazzetti, M. (2007, January 31). Iran may have trained attackers that killed 5 American soldiers, U.S. and Iraqis say. *The New York Times*. Retrieved October 21, 2008, from http://www.nytimes.com/2007/01/31/world/middleeast/31karbala.html

Griffiths, T. L., Steyvers, M., & Firl, A. (2007). Google and the mind: Predicting fluency with PageRank. *Psychological Science*, *18,* 1069–1076.

Gross, K., & D'Ambrosio, L. (2004). Framing emotional response. *Political Psychology*, *25,* 1–29.

Hamann, S. B., Ely, T. B., Grafton, S. T., & Kilts, C.D. (1999). Amygdala activity related to enhanced memory for pleasant and aversive stimuli. *Nature Neuroscience*, *2,* 289–294.

Hayes, J., Schimel, J., & Williams, T. J. (2008). Fighting death with death: The buffering effects of learning that worldview violators have died. *Psychological Science*, *19,* 501–507.

Healy, A. F., Aylward, A. G., Bourne, L. E., Jr., & Beer, F. A. (2009). Terrorism after 9/11: Reactions to simulated news reports. *American Journal of Psychology*, *122,* 153–165.

Heathcote, A., Raymond, F., & Dunn, J. (2006). Recollection and familiarity in recognition memory: Evidence from ROC curves. *Journal of Memory and Language*, *55,* 495–514.

Heuer, F., & Reisberg, D. (1990). Vivid memories of emotional events: The accuracy of remembered minutiae. *Memory & Cognition*, *18,* 496–506.

Jeffreys, H. (1998). *Theory of probability* (3rd ed.). Oxford, UK: Oxford University Press. (Original work published 1939).

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, and Cognition*, *20,* 1420–1436.

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

Jonathan, T., Alison, L., & Long, A. (2007). The influence of media coverage on responses to a hypothetical burglary scenario. *Journal of Applied Social Psychology*, *37,* 631–657.

Kassin, S. M., & Sukel, H. (1997). Coerced confessions and the jury: An experimental test of the "harmless error" rule. *Law and Human Behavior*, *21,* 27–46.

Kensinger, E. A. (2007). Negative emotion enhances memory accuracy: Behavioral and neuroimaging evidence. *Current Directions in Psychological Science*, *16,* 213–218.

Kensinger, E. A. (2008). Age differences in memory for arousing and nonarousing emotional words. *Journal of Gerontology Series B: Psychological Sciences and Social Sciences*, *63,* 13–18.

Kensinger, E. A. (2009). Remembering the details: Effects of emotion. *Emotion Review*, *1,* 99–113.

Kensinger, E. A., & Corkin, S. (2003a). Memory enhancement for emotional words: Are emotional words more vividly remembered than neutral words? *Memory & Cognition*, *31,* 1169–1180.

Kensinger, E. A., & Corkin, S. (2003b). Effect of negative emotional content on working memory and long-term memory. *Emotion*, *3,* 378–393.

Kensinger, E. A., & Schacter, D. L. (2006). Reality monitoring and memory distortion: Effects of negative, arousing content. *Memory & Cognition*, *34,* 251–260.

Killeen, P. R. (2005). An alternative to null-hypothesis significance tests. *Psychological Science*, *16,* 345–353.

Krueger, J. (2001). Null hypothesis significance testing: On the survival of a flawed method. *American Psychologist*, *56,* 16–26.

Kull, S., Ramsay, C., & Lewis, E. (2003). Misperceptions, the media, and the Iraq war. *Political Science Quarterly*, *118,* 569–598.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104,* 211–240.

Langeslag, S. J. E., & van Strien, J. W. (2008). Age differences in the emotional modulation of ERP old/new effects. *International Journal of Psychophysiology*, *70,* 105–114.

Levine, L. J., & Pizarro, D. A. (2004). Emotion and memory research: A grumpy overview. *Social Cognition*, *22,* 530–554.

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, F. Morgan, & J. Clare (Eds.), *Terrorism and torture: An interdisciplinary perspective*. Cambridge, UK: Cambridge University Press.

MacKay, D. G., Shafto, M., Taylor, J. K., Marian, D. E., Abrams, L., & Dyer, J. R. (2004). Relations between emotion, memory, and attention: Evidence from taboo Stroop, lexical decision, and immediate memory tasks. *Memory & Cognition*, *32,* 474–488.

Mather, M. (2007). Emotional arousal and memory binding: An object-based framework. *Perspectives on Psychological Science*, *2,* 33–52.

Mather, M., Mitchell, K. J., Raye, C. L., Novak, D. L., Greene, E. J., & Johnson, M. K. (2006). Emotional arousal can impair feature binding in working memory. *Journal of Cognitive Neuroscience*, *18,* 614–625.

McCloskey, M., Wible, C. G., & Cohen, N. J. (1988). Is there a special flashbulb-memory mechanism?

Journal of Experimental Psychology: General, *117,* 171–181.

McGaugh, J. L. (2000). Memory—a century of consolidation. *Science*, *287,* 248–251.

McNally, R. J., Clancy, S. A., Barrett, H. M., & Parker, H. A. (2004). Inhibiting retrieval of trauma cues in adults reporting histories of childhood sexual abuse. *Cognition & Emotion*, *18,* 479–493.

Neisser, U., & Harsch, N. (1992). Phantom flashbulbs: False recollections of hearing the news about Challenger. In E. Winograd & U. Neisser (Eds.), *Affect and accuracy in recall* (pp. 9–32). New York: Cambridge University Press.

Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior*, *32,* 303–330.

Oberauer, K., & Bialkova, S. (2009). Accessing information in working memory: Can the focus of attention grasp two elements at the same time? *Journal of Experimental Psychology: General*, *138,* 64–87.

Ost, J., Granhag, P.-A., Udell, J., & af Hjelmsäter, E. R. (2008). Familiarity breeds distortion: The effects of media exposure on false reports concerning media coverage of the terrorist attacks in London on 7 July 2005. *Memory*, *16,* 76–85.

Pezdek, K. (2003). Event memory and autobiographical memory for the events of September 11, 2001. *Applied Cognitive Psychology*, *17,* 1033–1045.

Phelps, E. A., & Sharot, T. (2008). How (and why) emotion enhances the subjective sense of recollection. *Current Directions in Psychological Science*, *17,* 147–152.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & the R Core team. (2009). nlme: Linear and nonlinear mixed effects models. R package Version 3.1–92 [Computer software]. Available from: http://CRAN.R-project.org/package=nlme

Porter, S., Taylor, K., & ten Brinke, L. (2008). Memory for media: Investigation of false memories for negatively and positively charged public events. *Memory*, *16,* 658–666.

R Development Core Team. (2008). R: A language and environment for statistical computing [Computer software]. Available from http://www.R-project.org

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25,* 111–196.

Rapp, D. N., & Kendeou, P. (2007). Revising what readers know: Updating text representations during narrative comprehension. *Memory & Cognition*, *35,* 2019–2032.

Ridley, A. M., & Clifford, B. R. (2004). The effects of anxious mood induction on suggestibility to misleading post-event information. *Applied Cognitive Psychology*, *18,* 233–244.

Rouder, J. N., Speckman, P. I., Sun, D., & Morey, R. D. (2009). Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, *16,* 225–237.

Schmidt, S. R. (2004). Autobiographical memories for the September 11th attacks: Reconstructive errors and emotional impairment of memory. *Memory & Cognition*, *32,* 443–454.

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.

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, CA: Academic Press.

Sharot, T., & Yonelinas, A. P. (2008). Differential time-dependent effects of emotion on recollective experience and memory for contextual information. *Cognition*, *106,* 538–547.

Sheridan, L. P. (2006). Islamophobia pre- and post-September 11th, 2001. *Journal of Interpersonal Violence*, *21,* 317–336.

Sinaceur, M., Heath, C., & Cole, S. (2005). Emotional and deliberative reactions to a public crisis: Mad cow disease in France. *Psychological Science*, *16,* 247–254.

Solomon, S., Greenberg, J., & Pyszczynski, T. (1991). A terror management theory of social behaviour: The functions of self-esteem and cultural worldviews. In M.P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 24, pp. 93–159). San Diego, CA: Academic Press.

Talarico, J. M., & Rubin, D. C. (2003). Confidence, not consistency, characterizes flashbulb memories. *Psychological Science*, *14,* 455–461.

Talarico, J. M., & Rubin, D. C. (2007). Flashbulb memories are special after all; in phenomenology, not accuracy. *Applied Cognitive Psychology*, *21,* 557–578.

Thompson, C. P., Skowronski, J. J., Larsen, S. F., & Betz, A. (1996). *Autobiographical memory: Remembering what and remembering when*. Mahwah, NJ: Lawrence Erlbaum Associates.

Tryon, W. W. (2001). Evaluating statistical difference, equivalence, and indeterminacy using inferential confidence intervals: An integrated alternative

method of conducting null hypothesis statistical tests. *Psychological Methods*, *6*, 371–386.

Unz, D., Schwab, F., & Winterhoff-Spurk, P. (2008). TV news—the daily horror? Emotional effects of violent television news. *Journal of Media Psychology*, *20,* 141–155.

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). Mahwah, NJ: Lawrence Erlbaum Associates.

Veno, A., & van den Eynde, J. (2007). Moral panic neutralization project: A media-based intervention. *Journal of Community and Applied Social Psychology*, *17,* 490–506.

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of *p* values. *Psychonomic Bulletin & Review*, *14,* 779–804.

Wagenmakers, E.-J., Lee, M. D., Lodewyckx, T., & Iverson, G. (2008). Bayesian versus frequentist inference. In H. Hoijtink, I. Klugkist, & P. A. Boelen (Eds.), *Bayesian evaluation of informative hypotheses* (pp. 181–207). New York: Springer.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54,* 1063–1070.

Wetzels, R., Raaijmakers, J. G. W., Jakab, E., & Wagenmakers, E.-J. (2009). How to quantify support for and against the null hypothesis: A flexible WinBUGS implementation of a default Bayesian *t* test. *Psychonomic Bulletin & Review*, *16,* 752–760.

Wilkes, A. L., & Leatherbarrow, M. (1988). Editing episodic memory following the identification of error. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *40,* 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: Human Experimental Psychology*, *52,* 165–183.

Windmann, S., & Kutas, M. (2001). Electrophysiological correlates of emotion-induced recognition bias. *Journal of Cognitive Neuroscience*, *13,* 577–592.

APPENDIX

Materials for all experiments

Experiment 1

Instructions study phase. Please read through the following report—consisting of a series of messages—as you would read a newspaper article. Once you have read a message, turn the page to read the next one, without back-tracking. You will be asked to recall the information later.

Series of messages

Low-emotionality condition

1. A plane departed from Sydney Airport at approximately 8.30 a.m. Eastern Standard Time. It carried 116 passengers, headed for Brisbane.
2. A plane crash was reported in the rural area of New South Wales by a nearby resident. The situation required immediate response from emergency services, the call being made at approximately 9.00 a.m. Eastern Standard Time.
3. Emergency services arrived 15 minutes later, working as fast as possible despite all of the damage to the plane, and debris everywhere. After a few hours of rummaging through the crash site, only 12 survivors had been found, 3 of whom were crew members.
4. [*Initial misinformation*]: The investigators on scene routinely went through possible causes of the crash. The media station reporting live from the crash site reported that bad weather was to blame for the crash.
5. Three crash site workers were taken to hospital as a precaution after some debris had fallen on them, but they seemed to only suffer minor injuries.
6. [*No-retraction*]: No Message 6, continued with Message 7.

6. [*Retraction-only*]: A few days later, investigation into the plane crash had been completed, and it was concluded that bad weather was not to blame.
6. [*Causal-alternative*]: A few days later, investigation into the plane crash had been completed, and it was concluded that bad weather was not to blame. The actual cause was determined to be a faulty fuel tank.
7. The number of casualties was confirmed, with only 15 survivors in total, 11 of whom were badly injured as a result of the crash. The investigation was closed exactly one week after the crash occurred.

High-emotionality condition (Messages 1–3, 5, & 7 identical to low-emotionality condition)

4. [*Initial misinformation*]: The investigators on scene routinely went through possible causes of the crash. The media station reporting live from the crash site reported that terrorists were to blame for the crash, claiming bomb fragments were found in the wreckage.
6. [*No-retraction*]: No Message 6, continued with Message 7.
6. [*Retraction-only*]: A few days later, investigation into the plane crash had been completed, and it was concluded that no foul play occurred, and terrorists were not to blame.
6. [*Causal-alternative*]: A few days later, investigation into the plane crash had been completed, and it was concluded that no foul play occurred, and terrorists were not to blame. The actual cause was determined to be a faulty fuel tank.

Alternative-throughout control group (Messages 1–3, 5, & 7 identical to low-emotionality condition, no Message 6)

4. [*Alternative causal referent—no initial misinformation, no retraction*]: The investigators on scene routinely went through possible causes of the crash. The media station reporting live from the crash site claimed that a faulty fuel tank was the cause of the crash.

Test questionnaire

Inference questions

1. Is there any evidence of careless management of the aircraft?
2. How could this crash have been avoided?
3. What precautions could be taken in the future to ensure this doesn't happen again?
4. What aspect of the plane crash should have been focussed on during investigation?
5. Does any aspect of the incident require further investigation?
6. Why would the media report such an assumption regarding the cause of the crash?
7. What would be a more plausible alternative cause for the crash than the media's report?
8. Why might the public be outraged by this incident?
9. What do you think was the true cause of the plane crash?

Experiment 2

Series of messages

Low-emotionality condition

1. A plane originating from Auckland (New Zealand) departed from Melbourne Tullamarine Airport at approximately 6.30 a.m. Western Standard Time. It carried 116 passengers, mainly New Zealanders and Australians, and was headed for Perth.
2. A plane crash was reported near the small town of Jolap Creek in rural Western Australia—600 km from Perth—by a nearby resident. The situation required immediate response from emergency services, the call being made at approximately 9.30 a.m. Western Standard Time.
3. Emergency services arrived 15 minutes later, working as fast as possible despite all of the damage to the plane, and debris everywhere. After a few hours of rummaging through the crash site, only 12 survivors had been found, 3 of whom were crew members.
4. [*Initial misinformation*]: The investigators on scene routinely went through possible causes of the crash. The media station reporting live from the crash site reported that bad weather was to blame for the crash.

5. Three crash site workers were taken to hospital as a precaution after some debris had fallen on them, but they seemed to only suffer minor injuries.
6. [*No-retraction*]: No Message 6, continued with Message 7.
6. [*Retraction-only*]: A few weeks later, investigation into the plane crash had been completed, and it was concluded that bad weather was not to blame.
6. [*Causal-alternative*]: A few weeks later, investigation into the plane crash had been completed, and it was concluded that bad weather was not to blame. The actual cause was determined to be a faulty fuel tank.
7. The number of casualties was confirmed, with only 12 survivors in total, 11 of whom were badly injured as a result of the crash. The investigation was closed exactly one month after the crash occurred.

High-emotionality condition (Messages 2, 3, 5, & 7 identical to low-emotionality condition)

1. A plane originating from Tehran (Iran) departed from Dubai International Airport at approximately 11.30 p.m. Western Standard Time. It carried 116 passengers, mainly Arabs and Australians, and was headed for Perth.
4. [*Initial misinformation*]: The investigators on scene routinely went through possible causes of the crash. The media station reporting live from the crash site reported that terrorists were to blame for the crash, claiming bomb fragments were found in the wreckage.
6. [*No-retraction*]: No Message 6, continued with Message 7.
6. [*Retraction-only*]: A few weeks later, investigation into the plane crash had been completed, and it was concluded that no foul play had occurred, and terrorists were not to blame.
6. [*Causal-alternative*]: A few weeks later, investigation into the plane crash had been completed, and it was concluded that no foul play had occurred, and terrorists were not to blame. The actual cause was determined to be a faulty fuel tank.

Alternative-throughout control group (Messages 2, 3, 5, & 7 identical to low-emotionality condition, no Message 6)

1. A plane departed from Melbourne Tullamarine Airport at approximately 6.30 a.m. Western Standard Time. It carried 116 passengers and was headed for Perth.
4. [*Alternative causal referent—no initial misinformation, no retraction*]: The investigators on scene routinely went through possible causes of the crash. The media station reporting live from the crash site claimed that a faulty fuel tank was the cause of the crash.

Test questionnaire

Additional inference questions not used in Experiment 1

6. Why were there so few survivors?
7. Why did it take so long for the investigations to complete?
10. Why does the media's account of the plane crash seem plausible enough to report?

Recall questions

1. What city was the plane headed for?
2. What was the original city of departure?
3. How many passengers were on the plane?
4. At what time was the phone call made to emergency services reporting the crash?
5. Where did the plane crash occur?
6. How many survivors were there?
7. What was initially reported by the media to be the cause of the crash?
8. How many crash site workers were taken to hospital?
9. Was this report challenged in any way?
10. How long did it take for the crash investigation to be completed?
11. What was determined to be the cause of the crash?
12. How long did it take emergency services to get to the scene of the crash?

Experiment 3

Instructions study phase. Dear participant, thank you for taking part in this study. On the following pages,

you will be presented with some edited excerpts of newspaper articles published in the Sydney Morning Herald in 1989. Please read through them as you would read any newspaper article. Once you have read a message, turn the page to read the next one, without back-tracking. You will be asked to recall some of the information later.

Series of messages

Low-emotionality condition. Plane fireball

1. A plane crash has occurred in a rural area of Northern France last night, near the town of Strasbourg, severely injuring two crew members, but leaving all passengers alive and well, including a 29-year-old Sydney husband and father-of-two. The incident occurred around 11 pm local time.
2. The British Caledonian Airways plane had departed from Ellinikon International Airport in Athens (Greece) at 7 pm. It carried 137 passengers and 11 crew, and was headed for London Heathrow.
3. It is believed the pilot had attempted an emergency landing on a remote crop field. Both the pilot and the co-pilot survived severely injured and were taken to a nearby hospital. Their condition has been described as serious but stable.
4. According to eye-witnesses, a massive explosion occurred only minutes after the emergency landing.
5. Police officials had earlier told TV that all passengers had managed to escape the aircraft before it burst into flames. Survivors had mainly suffered bruises and burn wounds.
6. [*Initial misinformation*]: The investigators on scene went through possible causes of the crash. The media station reporting live from the crash site reported that bad weather was to blame for the crash, claiming evidence was found that the plane had been struck by lightning, and citing survivors' reports of turbulences.
7. Three crash site workers suffering from respiratory problems were taken to hospital as a precaution, but were soon released.
8. [*No-retraction*]: Two months later, investigation into the plane crash had been completed.
8. [*Retraction-only*]: Two months later, investigation into the plane crash had been completed, and it was concluded that bad weather was not to blame.
8. [*Causal-alternative*]: Two months later, investigation into the plane crash had been completed, and it was concluded that bad weather was not to blame. The actual cause was determined to be a faulty fuel tank, which disrupted the plane's fuel supply.
9. In a press-conference, the carrier's spokesman offered support to the victims' families. He explained that crash investigations had been closed exactly 60 days after the crash had occurred.

Medium-emotionality condition (Messages 2–4, & 6–9 identical to low-emotionality condition)

1. A plane crash has occurred in a rural area of Northern France last night, near the town of Strasbourg, killing at least 100 people, including a 29-year-old Sydney husband and father-of-two. The incident occurred around 11 pm local time.
5. Police officials had earlier told TV that only few passengers had managed to escape the aircraft before it burst into flames. Survivors had mainly suffered bruises and burn wounds.

High-emotionality condition (Messages 1, 3–5, 7, 8—no-retraction—& 9 identical to medium-emotionality condition)

2. The British Caledonian Airways plane had departed from Mehrabad International Airport in Tehran (Iran) at 7 pm. It carried 137 passengers and 11 crew, and was headed for London Heathrow.
6. [*Initial misinformation*]: The investigators on scene went through possible causes of the crash. The media station reporting live from the crash site reported that terrorists were to blame for the crash, claiming bullet holes had been discovered in the cabin wreckage, and citing survivors' reports of an "Allahu Akbar" (Arabic for "God is great") scream when the plane went down.

8. [*Retraction-only*]: Two months later, investigation into the plane crash had been completed, and it was concluded that terrorists were not to blame.
8. [*Causal-alternative*]: Two months later, investigation into the plane crash had been completed, and it was concluded that terrorists were not to blame. The actual cause was determined to be a faulty fuel tank, which disrupted the plane's fuel supply.

Alternative-throughout control group (Messages 1–5, 7, & 9 identical to low-emotionality condition)

6. [*Alternative causal referent—no initial misinformation, no retraction*]: The investigators on scene went through possible causes of the crash. The media station reporting live from the crash site reported that a faulty fuel tank was to blame for the crash, citing a radio message from the pilot that had been received by an air-traffic controller at Paris airport minutes before the crash.
8. Two months later, investigation into the plane crash had been completed, and the actual cause was determined to be a faulty fuel tank, which disrupted the plane's fuel supply.

Test questionnaire

Replacement inference questions not used in Experiment 2

6. How would you explain the number of casualties?

Replacement recall questions not used in Experiment 2

16. At what time did the crash occur?
18. How many people were killed?
21. Was there any retraction or correction of this initial report?
24. How old was the Sydney husband and father-of-two?