

Running head: ATTITUDES AND MISINFORMATION

*****ARTICLE IN PRESS AT MEMORY & COGNITION*****

Do people keep believing because they want to? Pre-existing attitudes and the continued
influence of misinformation

Ullrich K. H. Ecker, Stephan Lewandowsky, Olivia Fenton, and Kelsey Martin

School of Psychology

University of Western Australia

Ullrich K. H. Ecker,

School of Psychology

University of Western Australia

Crawley, W.A. 6009, Australia

Telephone: +618 6488 3257

Fax: +618 6488 1006

E-mail: ullrich.ecker@uwa.edu.au

Abstract

Misinformation—defined as information that is initially assumed to be valid but is later corrected or retracted—often has an ongoing effect on people’s memory and reasoning. We tested the hypotheses that (a) reliance on misinformation is affected by people’s pre-existing attitudes, and that (b) attitudes determine the effectiveness of retractions. In two experiments, participants scoring higher and lower on a racial prejudice scale read a news report regarding a robbery. In one scenario, the suspects were initially presented as being Australian Aboriginals, whereas in a second scenario, a hero preventing the robbery was introduced as an Aboriginal person. These critical, race-related pieces of information were later retracted. We measured participants’ reliance on misinformation in response to inferential reasoning questions. Results showed that pre-existing attitudes influence people’s use of attitude-related information, but not the way in which a retraction of that information is processed.

Keywords: Misinformation; continued influence effect; attitudes; beliefs; motivated reasoning

Do people keep believing because they want to? Pre-existing attitudes and the continued
influence of misinformation

Misinformation—defined as information that is initially believed to be valid but which is subsequently retracted or corrected¹—has an ongoing impact on people’s memory and inferential reasoning even after unambiguous and clear retractions. For example, when people make inferences regarding the causal chain leading up to an event (e.g., the circumstances of a fire), misinformation (e.g., an initial suspicion of arson that is later corrected) is often relied upon even when people accurately remember its retraction (Ecker, Lewandowsky, & Apai, 2011a; Ecker, Lewandowsky, Swire, & Chang, 2011b; Ecker, Lewandowsky, & Tang, 2010a; H. Johnson & Seifert, 1994; Lewandowsky, Ecker, Seifert, Schwarz, & Cook, 2012; Wilkes & Leatherbarrow, 1988).

In these studies, participants typically read a news report about a fictional event, in which a piece of causal information is first given and then retracted for one group of participants. Participants are subsequently given a questionnaire asking them to make inferences about the event in response to indirect questions (e.g., in the present example, regarding the cause of the fire, or the response from authorities). References to the initial piece of misinformation are then tallied and compared to another group that did not receive a retraction. The typical result is that a retraction at most halves the number of references to a piece of misinformation, but does not eliminate its influence altogether (cf. Lewandowsky et al., 2012, for a review).

Previous research has offered some suggestions why this *continued influence effect* of misinformation (H. Johnson & Seifert, 1994) arises. Most of these theoretical accounts refer to failures of strategic memory processing. They argue that retracted or outdated information remains available in memory despite retractions or attempts to update memory (cf. Ayers &

Reder, 1998; Bjork & Bjork, 1996; Ecker, Lewandowsky, Oberauer, & Chee, 2010b; Kendeou & O'Brien, in press; Oberauer & Vockenberg, 2009). If this retracted but available information is automatically activated, it might be accepted as valid at face value; in particular, when its processing appears fluent, people might use a heuristic that fluency implies veracity (cf. Ecker, Swire, & Lewandowsky, in press; M. Johnson, Hashtroudi, & Lindsay, 1993; Lewandowsky et al., 2012; Schwarz, Sanna, Skurnik, & Yoon, 2007). Hence any automatic activation of outdated or invalidated information will require some strategic memory processing to counteract the potential impact of the automatically retrieved but invalid information. This strategic memory processing could involve the recollection of contextual details such as the source of the information or the details of the correction, or it could rely on a strategic monitoring process that determines the validity of an automatically retrieved piece of information (cf. Ecker et al., 2011b; Gilbert, Krull, & Malone, 1990; M. Johnson et al., 1993). The fact that explicit warnings about the misinformation effect reduce people's reliance on misinformation substantially (but not completely) has been taken as evidence for such a dual-process account of automatic and strategic retrieval processing (Ecker et al., 2010a).

Yet, this purely cognitive explanation does not consider motivational factors. A person processing information (including misinformation and retractions) is not a *tabula rasa*: people have pre-existing opinions and attitudes and process information in relation to what they already know and believe. Hence in many real-world circumstances, people will have a motivation to believe one event version over another, that is, people will often have an intrinsic motivation to resist a retraction.

Consequently, survey research suggests that in the real world, attitudes play a major role in how people process misinformation and retractions. For example, Kull, Ramsay, and Lewis

(2003) investigated misperceptions regarding the 2003 Iraq war, including the belief that weapons of mass destruction (WMD) had been found in Iraq after the invasion. During this period, there were many suspected WMD findings reported in the media, all of which were later retracted. Despite the extensive media coverage of the *failures* to find WMDs, a substantial proportion of the U.S. public continued to believe that WMDs had been found (see also Jacobson, 2010) and these people also showed much stronger support for the war than people who correctly believed that no WMDs had been discovered. Similarly, Travis (2010) reported that ongoing belief in the clearly refuted assertion that President Obama was born outside the U.S. was much more widespread among Republicans than Democrats. In a different context, a study carried out in the UK demonstrated that concerns about the refuted link between a common vaccine and autism (cf. Ratzan, 2010) co-varied with distrust in the public health system and the government's role in regulating health risks (Casiday, Cresswell, Wilson, & Panter-Brick, 2006). In Australia, Pedersen, Attwell, and Heveli (2005) reported that false beliefs about asylum seekers were predicted by political position and strength of national identity.

While these results demonstrate how attitudes determine (or at least co-vary with) people's beliefs in common misconceptions², they do not directly address a very important question: Do attitudes affect the processing of retractions directly? Surveys can shed light on people's belief in misinformation after its retraction, but surveys are rarely administered both before and after a retraction of misinformation. It therefore remains unclear whether an attitude-congruent belief after a retraction (a) simply mirrors the attitude-congruent belief *before* the retraction, or whether (b) it reflects the ineffectiveness of an attitude-incongruent retraction. For example, possibility (a) suggests a person mistrusting the public health system might have believed speculations about a vaccine-autism link more than a person trusting the public health

system, both before and after the retraction (with the retraction thus potentially having the same quantitative effect on both people). Alternatively, possibility (b) suggests that both people might have believed the initial suggestion to a similar degree, but the retraction may have reduced misbelief only in the trusting person.

Both these possibilities are plausible. Misinformation that supports one's attitudes will be consistent with existing personal knowledge and other beliefs, it will be familiar and therefore easy to process and more readily believed (cf. Dechêne, Stahl, Hansen, & Wänke, 2010; Schwarz et al., 2007), and it will often come from a trusted source and will be shared by others in one's social network. These factors may lead to attitude-dependent acceptance of misinformation (cf. Lewandowsky et al., 2012). On the other hand, retractions that contradict one's worldview will be less consistent with existing beliefs, less familiar, and will be more likely to come from an untrustworthy source and will not be shared by peers. These factors may undermine a retraction's effectiveness (cf. Lewandowsky et al., 2012).

Lewandowsky, Stritzke, Oberauer, and Morales (2005) found that people who were skeptical about the official motives for the 2003 Iraq war were better able to discount war-related misinformation more generally. In their study, participants read actual news items, some of which had been publicly retracted, and their belief in the items as well as memory for the retractions was measured. Lewandowsky et al. found that people who were skeptical about the official *casus belli* (i.e., they thought it was launched for reasons unrelated to WMD) showed reduced belief in pieces of information if they remembered their retraction, whereas memory for a retraction did not reduce belief in less skeptical participants. In other words, people who accepted the official reason for the war would continue to believe a retracted news story despite being able moments earlier to state explicitly that the story was false. This result suggests that

skepticism may be a mediating factor in the processing of retractions. Yet, arguably, people who were skeptical about the official reasons for the invasion of Iraq will have also had a different attitude regarding the war and war-related information (i.e., a more ‘anti-war’ attitude). This implies that to the degree that the retracted pieces of information used in the Lewandowsky et al. study were ‘pro-war,’ skeptical people would have been more likely to accept the retractions because they were more in line with their general ‘anti-war’ attitude. (All four “false-retracted” news items used in the Lewandowsky et al. study in fact portrayed the Allied forces as strong and successful, and/or the ‘enemy’ as weak, cruel, and fragmented; hence their retraction would have been in line with an ‘anti-war’ attitude.) Thus, the Lewandowsky et al. results could be taken to suggest that people are willing to accept retractions only to the degree that they are attitude-congruent.

Nyhan and Reifler (2010) addressed the issue in the field of political science. They presented Republicans and Democrats with a variety of political misperceptions (e.g., claims made by the George W. Bush administration that tax cuts in the early 2000’s had increased government revenue). Some of their participants were also given factual retractions of these misperceptions (e.g., a statement that government revenues had actually decreased as a result of the tax cuts). Retractions were found to be effective only when they were attitude-congruent. For example, Democrats relied less on the misinformation that revenue had increased after reading a retraction. In cases of attitude-incongruence, however, retractions actually backfired. That is, Republicans became even *more* likely to believe in the incorrect assertion of increased revenues after reading an attitude-incongruent retraction (this could be described as an extreme manifestation of option (b) discussed above). In a similar study, Nyhan, Reifler, and Ubel (2013) demonstrated that a correction of Sarah Palin’s “death panel” assertions was effective in

participants not supporting Palin, but backfired in Palin supporters (at least in those who were politically knowledgeable; also see Hart & Nisbet, 2012, who reported that Republicans became *less* supportive of climate mitigation policies when confronted with the potential health impacts of climate change). The occurrence of such extreme, backfiring effects is quite surprising, considering how counterintuitive it seems at first glance that people would modify their belief in a direction *counter* to evidence. Some experiments have found more subtle attitude effects, where corrections did not actually backfire, but where corrections were nonetheless relatively ineffective when they were attitude-incongruent (Nyhan & Reifler, 2010; also see Nyhan & Reifler, 2011). In sum, these findings suggest that attitudes have a major impact on the processing of retractions.

However, the results of Nyhan and Reifler (2010) and Nyhan et al. (2013) stand in contrast to some data from our own lab. In a re-analysis of data reported in Ecker et al. (2011a)—a study designed to investigate the effect of emotiveness on misinformation processing, using a plane crash scenario—we found that participants high in islamophobia relied more on the misinformation that a terrorist attack had caused the plane crash (compared to participants with low islamophobia scores). In contrast to the results of Nyhan and colleagues, however, we found that this group difference was present and of comparable magnitude both before and after a retraction. In other words, the retraction had the same effect independent of its attitude congruence, reducing the number of references to a terrorist attack in both groups to an equal extent (this is a manifestation of option (a) discussed above). This finding must be considered provisional because it was based on a post-hoc analysis of data collected for a different purpose, available only for a subsample of participants. Nonetheless, other researchers have also failed to find support for attitude-driven backfire effects (Berinsky, 2012; Garrett,

Nisbet, & Lynch, in press). For example, Berinsky reported that corrections of the rumor that U.S. health care changes would promote euthanasia were effective in both Democrats and Republicans.

The present study was designed to further investigate the interplay of attitudes and the processing of misinformation and retractions. Participants scoring higher and lower on a measure of racial prejudice read fictional accounts of crimes supposedly involving indigenous Australians; this aspect of the events—that is, the race of the protagonists—was subsequently retracted for one group of participants. Continued reliance on retracted misinformation during a subsequent inference task was then compared to the responses in another group that received no retraction. We expected the racial prejudice factor to have a substantial effect on the overall number of references to the critical race-related information. We also expected a retraction to significantly reduce the number of references to the critical information. We had no strong expectation regarding the question of whether racial attitudes would have an impact on the effectiveness of the retraction (i.e., the reduction in references to the critical information caused by the retraction), given the contradictory results obtained in previous research. Unlike most of the previous (survey-based) research, our experiment was designed to shed light on this issue as it featured both non-retraction and retraction conditions.

Experiment 1

Method

In Experiment 1, participants scoring higher vs. lower on a racial prejudice questionnaire were presented with a fictitious news report about a liquor store robbery. There were three versions of the report, which differed in the description of the suspects. Two versions initially described the suspects as Aboriginal Australians; this information was later retracted in one

version of the report (the retraction condition) but not in the other (the no-retraction control condition). The third version (the no-misinformation control condition) described the suspects as Caucasian and also contained no retraction. The two control conditions provided a ceiling and baseline, respectively, against which to assess the effects of the retraction. The no-misinformation condition was included because any description of a liquor store robbery in Australia may lead some people to assume that the suspects were Aboriginal based on common stereotypes (cf. Wimshurst, Marchetti, & Allard, 2004); we thus expected this condition to yield a non-zero baseline. The experiment employed a 2 (racial prejudice: high vs. low) \times 3 (retraction condition: no-misinformation, no-retraction, retraction) between-subjects design.

The racial prejudice scale used was the Attitudes Toward Indigenous Australians scale (ATIA; Pedersen, Beven, Walker & Griffiths, 2004), an 18-item questionnaire with good validity and reliability (reported internal consistency $\alpha = .93$; Pedersen et al., 2004), measuring racial prejudice towards indigenous Australians on 7-point Likert scales.

Participants. A priori power analysis (G*Power 3; Faul, Erdfelder, Lang, & Buchner, 2007) suggested that in order to detect a medium-size effect of $\eta_p^2 = .1$ at $\alpha = .05$ and $1-\beta = .80$, the minimum sample size should be at least 90. In line with precedents (see in particular Ecker et al., 2011a), we decided to test a total of $N = 144$ participants, all undergraduate students from the University of Western Australia (97 females, 47 males; age range 17-46; mean age 19 years).³ Approximately a third of these participants ($n = 47$) were randomly sampled from the upper and lower quartiles of a population of students pre-screened with the ATIA ($N = 379$; internal consistency of the ATIA in this population was $\alpha = .91$). The remaining two thirds were not pre-screened for pragmatic reasons (i.e., lack of access to a pre-screened population). All participants completed the ATIA (again) during the experimental session (test-retest reliability on the

subsample that completed the ATIA twice was high, $\rho = .90$). Participants were randomly assigned to the different retraction conditions, and divided into higher and lower racial prejudice groups based on median splits on their (more recent) ATIA score ($n = 24$ per cell).

Stimuli. Participants were given a fictitious news report, consisting of a series of 14 messages, each printed on a separate page, which provided an account of a liquor store robbery in Australia's Northern Territory. Across conditions, the stories differed only at message 5, when the critical information about the race of the suspects was introduced, and message 11, where that piece of critical information was either retracted or not (see Appendix for materials).

In the no-misinformation condition, message 5 stated that "*police...believed the three suspects were Caucasian*" and message 11 gave the neutral piece of repeated information that "*Police...confirmed that the owner of the store was the sole person in the store.*" In the retraction and no-retraction conditions, message 5 stated that "*police...believed the three suspects were Aboriginal.*" In the retraction condition, message 11 then stated that "*Police...no-longer believed the suspects were...Aboriginal.*" Message 11 in the no-retraction condition was identical to the neutral message 11 of the no-misinformation condition.

Participants' understanding of the story, and in particular their reliance on misinformation, was assessed using an open-ended questionnaire. The questionnaire contained 10 inference questions, 9 fact-recall questions, and 2 retraction-awareness questions (always given in this order to prevent any impact from fact retrieval on people's inferences; see Appendix).

Inference questions required participants to infer something about the circumstances surrounding the incident, and were designed to elicit responses indirectly related to the critical information, that is, the race of the suspects. For example, the inference question "*Why did the*

shop owner have difficulty understanding the conversation between the attackers?” could be answered by relying on the critical information (e.g., that the intruders were speaking in their native Aboriginal language), although there were other possible explanations (e.g., the attackers had their mouths covered or were intoxicated, and so on). The final inference question (“*Who do you think the attackers were?*”) was placed at the end of the fact-recall questions so it appeared to be a recall question, but it was scored as an inference question.

The nine fact-recall questions were used to assess participants’ memory for the factual details of the story (e.g., “*What sort of car was found abandoned?*”). These questions did not relate to the race of the suspects. Finally, the two retraction-awareness questions tested participants’ awareness of the retraction (e.g., “*Was any of the information in the story subsequently corrected or altered?*”). The fact-recall and retraction-awareness questions were included to control for insufficient encoding, allowing for the potential exclusion of participants who did not recall the event sufficiently well, or who may not have noticed the retraction at all.

Analyses thus focused on (a) the accuracy of recall (fact-recall score), (b) memory for the retraction (retraction-awareness score), and most important, (c) reliance on the critical information (inference score).

Procedure. Participants read the report at their own pace without backtracking; they were informed that their memory for and understanding of the report would be tested (to ensure adequate encoding). Reading was followed by a 30-minute retention interval that was filled with an unrelated memory updating computer task. Participants were then given the open-ended questionnaire and instructed to answer all questions in the order given. Finally, participants were given a bundle of three questionnaires to complete, the last of which was the ATIA. The ATIA was administered at the end of the procedure so as not to prime Aboriginal-related responses on

the open-ended questionnaire; the two other questionnaires were unrelated to the study. The entire experiment took approximately one hour.

Results

Racial prejudice scores. ATIA scores ranged from 0 to 4.56 (maximum possible score was 6). Across experimental conditions, the mean racial prejudice scores were 1.33 ($SE = 0.07$) and 2.91 ($SE = 0.08$) in the lower and higher racial-prejudice groups, respectively. We acknowledge that the ATIA score of the ‘high racial-prejudice’ group in this experiment was only moderate.⁴ The group difference was nonetheless significant, $F(1,142) = 220.63$, $MSE = 0.41$; $p < .001$, $\eta_p^2 = .61$. Mean racial prejudice scores for all cells are given in the upper panel of Table 1.

Questionnaire coding. All open-ended questionnaires were scored by a trained scorer who was blind to the experimental conditions, following a standardized guide. Another trained scorer scored 5 questionnaires from each condition to assess inter-rater reliability, which was found to be very high ($r > .95$ for inference, fact-recall, and retraction-awareness questions).

Fact-recall questions were scored 1 for correct responses and 0 for incorrect responses. For certain questions it was possible to receive partial marks of 0.5 or 0.33 for partially correct responses, as determined a priori in the scoring guide. As there were 9 fact-recall questions, the maximum fact-recall score was 9. Retraction-awareness questions were given a score of 1 if participants remembered the retraction and a score of 0 if they did not. The maximum retraction-awareness score was 2.

Inference questions were scored 0 or 1. Any uncontroverted mention of Aboriginal people, Aboriginal culture or communities, or anything that directly implied Aboriginal persons as the suspects of the robbery, was counted as a reference to the critical information and given an

inference score of 1. Examples of inferences scoring 1 would be the response “*The robbers were speaking in their Aboriginal language*” to the inference question “*Why did the shop owner have difficulty understanding the [...] the attackers?*”, or the response “*Because police rarely solve crimes involving Aboriginals*” to the inference question “*Why do the police fear the case will remain unsolved?*” In contrast, the response “*First they thought the robbers were Aboriginals, but apparently not*” to the inference question “*Who do you think the attackers were?*” would have been considered a controverted statement and given a score of 0. The maximum inference score was 10.

Accuracy of recall. Mean fact-recall accuracy rates (out of a maximum of 1) for high- and low-prejudice groups are given in the upper panel of Table 2. A two-way analysis of variance (ANOVA) with the factors retraction condition and racial prejudice revealed no significant effects, F 's < 2.84 , $p > .05$.

Awareness of retraction. Mean rates of retraction awareness (in the retraction condition; out of a maximum of 1) were .79 ($SE = 0.08$) and .85 ($SE = 0.06$) for the high- and low-prejudice groups, respectively. This difference was not significant, $F < 1$.

Inferential reasoning. The mean inference scores for both high- and low-prejudice groups across all conditions are shown in Figure 1.

Not surprisingly, the number of references to the critical information was lowest (but not zero; $M = 0.40$, $SE = 0.08$, $t(47) = 4.86$; $p < .001$) in the no-misinformation condition, when the critical piece of information was never explicitly given. Also as expected, the number of references to the critical information was highest in the no-retraction condition, when this information was introduced but never challenged.

We ran a two-way between-subjects ANOVA on mean inference scores, with the factors retraction condition and racial prejudice. The analysis revealed reliable main effects of retraction condition, $F(2,138) = 43.81$, $MSE = 1.35$; $p < .001$, $\eta_p^2 = .39$, and racial prejudice, $F(1,138) = 6.67$, $MSE = 1.35$; $p = .01$, $\eta_p^2 = .05$. The interaction between retraction condition and racial prejudice was not significant, $F(2,138) = 1.58$, $p = .21$. These effects were confirmed in an ANOVA excluding the no-misinformation control condition, yielding significant effects of retraction type, $F(1,92) = 16.21$, $MSE = 1.87$; $p < .001$, $\eta_p^2 = .15$, and racial prejudice, $F(1,92) = 4.36$, $MSE = 1.87$; $p = .04$, $\eta_p^2 = .05$, but no significant interaction, $F(1,92) = 2.01$, $p = .16$.

A number of planned contrasts, again with mean inference scores as the dependent variable, were conducted to assess the overall effectiveness of the retraction and to further examine the relationship between retraction type and racial prejudice. Results are reported in Table 3. First, we assessed the difference between the no-retraction and retraction conditions: the retraction significantly reduced the number of references to the critical information in both racial prejudice groups (Contrasts 1 and 2). We then investigated whether there was a significant reliance on misinformation after a retraction (i.e., a continued influence effect), by contrasting the retraction condition with the no-misinformation baseline condition. Contrasts 3 and 4 show significant continued reliance on misinformation in both racial prejudice groups despite the retraction. Finally, the effects of racial prejudice were investigated, and were shown only to be significant in the No-Retraction condition (Contrasts 5-7), meaning that people scoring relatively high on racial prejudice mentioned the race of the suspects more often than people in the low-prejudice group, but mainly when this information was explicitly supplied without being subsequently retracted.⁵

Discussion

Experiment 1 examined the relationship between racial attitudes and the continued influence of racial misinformation. In line with previous research, Experiment 1 found that a simple retraction significantly reduced reliance on misinformation, but did not eliminate reliance on misinformation (cf. Ecker et al., 2010a). This means that even participants in the low-prejudice group failed to fully discount the Aboriginal misinformation after a retraction; they continued to make a significant number of references to the misinformation. This suggests that strategic memory processes failed to suppress automatic activation of race-related misinformation, even in people who arguably are not predisposed towards maintaining a belief in the misinformation (i.e., Aboriginal robbers).

In terms of racial prejudice effects, we found that people with relatively high levels of racial prejudice made more references to attitude-congruent racial information, as long as this information was explicitly given and not retracted. However, relatively high racial prejudice did not lead to a failure to discount attitude-incongruent misinformation. On the contrary, the retraction of racial misinformation was equally effective in both prejudice groups. This finding is in line with the re-analysis of Ecker et al. (2011a) data reported at the outset, and it suggests that racial attitudes do not influence how people process a retraction of racial misinformation. If anything, the present retraction tended to be numerically more effective in the high-prejudice group (i.e., when attitude-incongruent), resulting in both groups making an equivalent number of references to Aboriginal misinformation after a retraction. In sum, the results provide evidence against the notion that people generally seek to reinforce their pre-existing attitudes by dismissing the retraction of misinformation.

One factor that may have influenced these results—in particular the lack of interaction between racial prejudice and the effectiveness of the retraction—is the use of a stereotypical scenario. The liquor-store scenario in Experiment 1 was congruent with negative stereotypes about Aboriginal people (Wimshurst et al., 2004), and strong innuendo (setting the incident in Australia’s Northern Territory, a region with a large indigenous population, the use of Aboriginal place names, etc.) was intentionally used to boost the number of references to the critical information in an attempt to avoid floor effects.⁶ As knowledge and use of stereotypes can be largely independent from people’s attitudes (Devine & Elliot, 1995), participants in the low-prejudice group might have referred to the misinformation despite a retraction based purely on the strong innuendo inherent in the story. This may have artificially inflated the level of post-retraction misinformation reliance in the low-prejudice group, potentially masking an interaction involving level of prejudice.

Moreover, the racial information may have been more salient for people relatively high in racial prejudice, and this group’s higher scores may simply reflect this difference in salience, rather than being directly related to their racial attitudes. Experiment 2 was designed to address these concerns by using a scenario that was incongruent with stereotypes.

Experiment 2

Experiment 2 was similar to Experiment 1, but the piece of misinformation now related to an Aboriginal hero that prevented a robbery, thus running counter to common stereotypes about Indigenous Australians.

Method

Participants scoring high vs. low on the ATIA racial prejudice questionnaire were presented with a news report about a bank robbery. There were two versions of the report. Both

initially described a citizen preventing the robbery as Aboriginal; this information was later retracted in one version of the report (the retraction condition) but not in the other (the no-retraction control condition). In contrast to Experiment 1, we omitted a third, no-misinformation control version as we did not expect above-zero scores in such a condition with the counter-stereotypical material used in this study. The no-retraction control condition therefore provided a ceiling level of references to the critical information, while the baseline was effectively zero. The experiment employed a 2 (racial prejudice: high vs. low) \times 2 (retraction condition: no-retraction, retraction) between-subjects design.

Participants. A priori power analysis (G*Power 3; Faul et al., 2007) suggested that in order to detect a medium-size effect of $\eta_p^2 = .1$ at $\alpha = .05$ and $1 - \beta = .80$, the minimum sample size should be 76. We tested a total of $N = 100$ undergraduate students from the University of Western Australia (69 females, 31 males; age range 17-36; mean age 19 years). Participants were sampled from a population of students pre-screened with the ATIA ($N = 728$).⁷ Participants were divided into high and low racial prejudice groups based on their ATIA score, and randomly assigned to the different retraction conditions ($n = 25$ per cell).

Stimuli. Participants were given a report describing a fictitious attempted bank robbery, which was prevented by an Aboriginal man who disarmed the perpetrator. The report was presented as a series of 14 messages via a Microsoft PowerPoint presentation (see Appendix). Each message was shown separately and displayed for a set amount of time (0.4 seconds per word). The reading time was derived from a pilot study on a different sample of undergraduate students ($N = 10$), and was calculated as the mean reading time plus 3 *SDs*; this allowed for comfortable reading while not providing excess slack time.

In both conditions, message 5 stated that there was “*a local Aboriginal resident who stepped in front of the main tellers and convinced the man [i.e., the intruder] to put down his gun.*” The two versions of the story only differed at message 11. In the retraction condition, message 11 stated that “*Police released a second statement revealing that the man who helped apprehend the intruder was not Aboriginal as was first reported.*” In the no-retraction condition, message 11 only repeated information given earlier: “*Police released a second statement [...] confirming that the intruder was carrying a gun.*”

Participants’ reliance on misinformation and their understanding of the story were assessed using a questionnaire containing 10 inference questions, 10 fact-recall questions, and two retraction-awareness questions (see Appendix). Again, inference questions were designed to elicit race-related responses, while also allowing for responses unrelated to the protagonist’s race. For example, the inference question “*Why did police say the course of events was ‘unexpected’?*” could be answered by referring to the Aboriginal man (e.g., because an Aboriginal hero may have been unexpected), or it could be answered without referring to the Aboriginal man (e.g., because not many people are brave enough to stand up to someone with a gun). The 10 fact-recall questions had no relation to the race of the protagonist (e.g., “*On which day did the incident occur?*”). The retraction-awareness questions were identical to Experiment 1.

Procedure. The procedure was identical to Experiment 1, with the three exceptions that (a) all participants were pre-screened with the ATIA, (b) participants read the report individually on a computer screen, not on paper, and (c) that encoding time was fixed.

Results

Racial prejudice scores. ATIA scores ranged from 0.11 to 6 (i.e., nearly the full range of possible scores, 0-6). Mean racial prejudice scores were 1.54 ($SE = 0.10$) and 4.22 ($SE = 0.15$) in

the low and high racial-prejudice groups, respectively. This was a significant difference, $F(1,98) = 218.10$, $MSE = 0.82$; $p < .001$, $\eta_p^2 = .91$. The cell means are given in the lower panel of Table 1.

Questionnaire coding. Scoring was identical to Experiment 1. Again, inter-rater reliability was found to be very high ($r > .93$ for inference, fact-recall, and retraction-awareness questions).

Accuracy of recall. Mean rates of accuracy for the fact-recall responses (out of a maximum of 1) are given in the lower panel of Table 2. A two-way analysis of variance (ANOVA) with the factors retraction condition and racial prejudice revealed no significant effects, F 's < 2.52 , p 's $> .10$.

Awareness of retraction. Mean rates of retraction-awareness (out of a maximum of 1) were .58 ($SE = 0.07$) and .68 ($SE = 0.07$) for the high- and low-prejudice groups, respectively. This was not a significant difference, $F < 1$.

Inferential reasoning. The mean inference scores for both high- and low-prejudice groups across retraction conditions are shown in Figure 2. As in Experiment 1, a retraction substantially reduced references to misinformation, and inference scores were higher when the critical information was attitude-congruent (i.e., in the low-prejudice group). In other words, a reversal of the stereotype also led to a reversal of the prejudice effect on inference scores.

We ran a two-way between-subjects ANOVA with the factors retraction condition and racial prejudice, which revealed reliable main effects of retraction condition, $F(1,96) = 12.13$, $MSE = 1.67$; $p = .001$, $\eta_p^2 = .11$, and racial prejudice, $F(1,96) = 6.52$, $MSE = 1.67$; $p = .01$, $\eta_p^2 = .06$. The interaction was not significant, $F < 1$. Results indicate that the retraction reduced the number of references to the critical information in both racial prejudice groups, and that

people in the low-prejudice group mentioned the race of the Aboriginal hero more often than people in the high-prejudice group, both before and after a retraction.

While the number of post-retraction references to misinformation was clearly no different from zero in the high-prejudice group ($t < 1$), the number was substantially above zero in the low-prejudice group, $M = 0.60$, $SE = 0.22$, $t(25) = 2.78$, $p = .01$. The number of references to the critical information was also above-zero in the no-retraction condition of the high-prejudice group, $M = 0.84$, $SE = 0.31$, $t(25) = 2.72$, $p = .01$.⁸

Discussion

Experiment 2 examined the relationship between racial attitudes and the continued influence of racial misinformation, using a stereotype-incongruent scenario. As in Experiment 1, we again found that a simple retraction significantly reduced reliance on misinformation. While a retraction did not eliminate reliance on misinformation in the low-prejudice group, the retraction did virtually eliminate misinformation effects in the high-prejudice group, meaning we found no continued influence effect in that group. In fact, only one person made a single reference to an Aboriginal ‘hero’ in the high-prejudice group after a retraction (fittingly, this was accompanied by the assumption that the Aboriginal man was cooperating with the robber, an assertion that a number of participants in the high-prejudice group made).

Concerning the effects of racial prejudice, our analyses demonstrated that people scoring low in racial prejudice referred to an Aboriginal hero more often, in particular when there was no retraction. As in Experiment 1, this shows that there were more references to the critical information when this information was attitude congruent. Since the scenario used in Experiment 2 did not conform to any relevant stereotypes, it seems very unlikely that stereotype-based responding had any impact on these results. Moreover, we argue that information

regarding an Aboriginal hero would be more salient to a racially prejudiced person. It follows that the present effects, with higher scores in the low-prejudice group, cannot be explained by salience.

We do not believe that the obvious floor effect is reason for concern. First of all, we did not find a significant interaction between retraction and racial prejudice despite the reduced variance associated with the floor effect. Secondly, the low range of inference scores was expected given the counter-stereotypical scenario. In particular, we expected the number of references to the critical information to go down to (almost) zero in the high prejudice group. Our primary interest lies thus in the substantial reduction of inferences in the low-prejudice group, which was unaffected by a floor effect and numerically larger (about 60 %) than the reduction expected based on the results of Experiment 1 (where a retraction decreased references to the critical information only by about 50 % in the comparable high-prejudice group, i.e., the group for which the retraction was attitude-incongruent).

General Discussion

In two experiments, we found that people use race-related information in their inferential reasoning mainly when this information is congruent with their attitudes. People scoring relatively high on a racial prejudice measure mentioned an Aboriginal crime suspect more often than people with low prejudice scores, while people with low prejudice scores mentioned an Aboriginal hero more often. The fact that the effects reversed when the scenario changed from a negative to a positive depiction of the Aboriginal person(s) is strong evidence that the usage of racial information was determined by people's attitudes, and not by other factors such as salience or reliance on stereotypes. This result provides experimental confirmation of survey-based research that has shown attitudes to be a major determinant of the information people believe and

use in their reasoning, be it the use of misinformation despite retractions (Casiday et al., 2006; Kull et al., 2003; Pedersen et al., 2005; Travis, 2010), or the (non-)belief in empirical evidence and (non-)support for actions based on empirical evidence (e.g., Aldy, Kotchen, & Leiserowitz, 2012; Fielding, Head, Laffan, Western, & Hoegh-Guldberg, 2012).

In contrast, people's racial attitudes did not determine the effectiveness of retractions. Retractions reduced reliance on the critical information, but they did so equally for people in the high- and low-prejudice groups. In other words, the differences between prejudice groups were roughly equivalent across retraction and no-retraction groups. It is thus possible that any group differences found in surveys measuring belief in retracted misinformation may simply reflect pre-retraction belief differences.

Our results contrast with findings by Nyhan and Reifler (2010) and Nyhan et al. (2013), who reported that retractions were effective mainly when they were attitude-congruent, and could even backfire when they were attitude-incongruent, which is in line with research on motivated reasoning. Motivated reasoning is biased information processing that serves to confirm pre-existing beliefs rather than objectively assessing the available evidence (for reviews of motivated cognition, see Kunda, 1990; Redlawsk, 2002; for a review from a misinformation perspective, see Ecker et al., in press). For example, Lord, Ross, and Lepper (1979) studied people who had strong opposing views on capital punishment, and presented both groups with two fictional articles: one supporting and one refuting the claim that capital punishment reduces crime. Each group was more convinced by the article that supported their own beliefs, and after reading and discussing both articles, the two groups differed even more in their opposing views than before. Redlawsk (2002) and Redlawsk, Civetinni, and Emmerson (2010) demonstrated how, at least up to a certain 'tipping point,' voters can increase their support for their favored

political candidates when faced with negative information about them. The most common explanation for such behavior is that motivated reasoners counter-argue attitude-incongruent information, in the process activating many arguments supporting their existing attitude ('attitude bolstering'; Jacks & Cameron, 2003; Prasad et al., 2009).

What reasons might underlie this discrepancy between our results and other results that show a strong effect of attitudes on the processing of discounting information? One obvious methodological difference is that our study used fictional materials, whereas most of the studies discussed above used real-world materials. We suggest this factor is unlikely to explain the observed differences on its own, given that Lord et al. (1979) used fictitious materials but nonetheless observed strong attitude effects. Arguably, as long as the fictional material is realistic and plausible, participants process it in a very similar manner to 'real-world' information (cf. Kreitman, 2006; Young, 2010).

Another difference between the two sets of studies is that our scenario involved a singular episodic event (i.e., a crime), whereas most of the other research looked at belief in more general assertions (e.g., whether capital punishment deters crime; Lord et al., 1979). Again, this factor by itself is unlikely to explain the differences in outcome because attitude effects have also been found with singular episodic material (e.g., whether President Bush misrepresented the effects of tax cuts; Nyhan & Reifler, 2010; see also Lewandowsky et al., 2005).

We thus propose three reasons for the discrepancy between our results and previous research reporting effects of attitudes on retraction processing: two related to the arguments of fictionality and singularity discussed above, and one related to the measurement of beliefs. First, it seems plausible to assume that real-world information is often encoded repeatedly before it is retracted, and may therefore require stronger retractions than information encoded only once (cf.

Ecker et al., 2011b). As people arguably seek out information that is in line with their attitudes (cf. T. Johnson, Bichard, & Zhang, 2009; Kunda, 1990), they may encode attitude-congruent real-world myths more often than fictional misinformation presented in the lab.

Second, we suggest that attitude-incongruent retractions will be effective to the degree that they do not require attitude change. To illustrate, imagine a situation where accepting a retraction will require attitude change: For a Republican supporting President Bush's tax cuts to acknowledge that Bush's claims about the tax cuts were incorrect (Nyhan & Reifler, 2010), it would require the person to accept that (a) a supported politician told an untruth, and that (b) it was a poor decision to support the associated policies. Those acknowledgments, in turn, would require a change in attitude regarding the tax cuts. Similarly, for a misinformed person, accepting evidence that goes against their misinformed belief—for example, evidence that vaccines do not cause autism—would inevitably require a shift in attitudes. There will thus be a strong resistance to an attitude-incongruent retraction if accepting it would inevitably induce attitudinal change. Thus, there will be continued reliance on misinformation whose correction would threaten one's worldview (Nyhan & Reifler, 2010, 2011).

In contrast, accepting that a particular crime was not committed by an Aboriginal person, or that a brave act was not performed by an Aboriginal, could be accepted without any change in attitude. In the case of the robbery in Experiment 1, one could still believe that most liquor store robberies in Australia are perpetrated by Aboriginal people, and that most Aboriginal people are criminals. In fact, a single robbery committed by a non-Aboriginal person does not constitute any evidence against these beliefs. In the case of the attempted bank robbery in Experiment 2, one could still believe that most Aboriginal people are brave and fearless despite the retraction that the hero was not an Aboriginal person; accepting the retraction does not constitute evidence

against this belief. Hence, if accepting a retraction does not require a shift in attitudes, it will seemingly be followed even when it is attitude-incongruent.

Attitude-incongruent retractions may also be effective when people can use strategies to avoid attitude change. For example, people can accommodate exceptions to stereotypes and thus maintain them by way of a process known as stereotype subtyping (Kunda & Oleson, 1995; Richards & Hewstone, 2001). This means, for example, that people with high racial prejudice might be able to accept an Aboriginal hero while maintaining their negative stereotype regarding Aboriginal people if they can identify a seemingly atypical attribute and use it to subtype the ‘deviant exemplar’ (e.g., they might argue that Aboriginal people are usually criminal and coward, but that ‘true’ Aboriginal people from ‘Outback’ Australia may be braver; cf. Pedersen et al., 2004).

In the absence of such an atypical attribute, it seems the only group in the present experiments that was highly motivated to believe one event version over the other was the high-prejudice group of Experiment 2. Participants in this group may have struggled to accept an Aboriginal hero, hence they were motivated to accept the attitude-congruent retraction. In fact, this was the first group of participants across all of our previous studies that did not show a continued influence effect (i.e., they showed an elimination of misinformation effects; although one has to consider the low overall rate of inferences). This group also showed indications of motivated reasoning, with some participants rationalizing that the Aboriginal man might have been an accomplice of the robber. Assuming that misinformation effects arise from a failure of strategic memory processes (cf. Introduction), this interpretation implies that an attitude-based motivation to believe one event version over another can lead to a boost in strategic

monitoring—in this case a high-prejudice group making sure to correct the initial attitude-incongruent event representation.

The third reason for the discrepancy between our results and other results suggesting attitude effects on the processing of retractions may lie in the difficulties that exist in the direct measurement of beliefs (e.g., in surveys). Clearly, what people *say* they believe and what they *actually* believe may be two different ‘animals,’ in particular in non-intimate social interactions (cf. Fazio & Olson, 2003; Lamont, 2007). This means that when people holding a certain belief are presented with convincing belief-incongruent evidence (such as a detailed refutation) they might change their belief in direction of the evidence, but they might not overtly acknowledge this change. This could be in an attempt to ‘save face,’ or it could be to instill doubt in the person presenting the evidence as to how convincing the evidence is.

Typically, and as briefly discussed earlier, backfire effects such as those reported by Nyhan and Reifler (2010) and Nyhan et al. (2013) are explained in terms of counter-arguing, attitude bolstering, and other motivated reasoning processes. Common to those explanations is the idea that people’s beliefs change in a direction counter to the presented evidence. An alternative view on backfire effects would hold that rather than representing attitude change, they reflect people’s attempt to defend and maintain their attitudes and beliefs. On this view, backfire effects are an occasional and inadvertent consequence of an overzealous attempt to maintain an attitude and protect it against change. Some support for this alternative explanation comes from a study by Gal and Rucker (2010). These authors argued that even though people often express their beliefs more vigorously following disconfirming evidence, this may be an ironic effect of reduced confidence in those beliefs. Specifically, people may have reduced their level of belief in line with the evidence, but the ensuing reduction in confidence in that belief may be threatening

to one's self-concept. In consequence, increased overt advocacy of the belief may be required to fend off this threat. Consonant with this idea, Gal and Rucker found that participants engaged in stronger belief advocacy when they were less confident of their beliefs.

To illustrate, in social interactions both empirical arguments and the expression of one's belief can be used in an effort to convince others of one's attitude or defend one's attitudes against persuasion; in this sense, attitudes are based on evidence and/or beliefs. This implies that when people (secretly) acknowledge that the real-world evidence landscape has shifted against their belief, they might adjust their belief, and potentially the underlying attitude, in direction of the evidence. Yet, at the same time, they might move their *expressed* belief in the opposite direction, in order to maintain a perceived balance of evidence and expressed belief in favor of their initial attitude (also see Batson, 1975, for a similar argument). As long as a positive (i.e., attitude-congruent) balance can be maintained, that is, as long as negative (i.e., attitude-incongruent) evidence can be counterbalanced by positive increments in expressed belief, the attitude can be defended and (ostensibly) maintained. If the negative evidence becomes overwhelming, however, there might be a 'tipping point' at which the expressed belief cannot sensibly be increased any more. At this point, attitude change may occur and may also be acknowledged (cf. Redlawsk et al., 2010).

Conclusion

In two experiments, we showed that pre-existing attitudes co-determine people's reliance on (mis-)information. That is, people are more likely to use a piece of information in their reasoning when this piece of information is congruent with their attitudes and beliefs. Unlike some previous research, however, we found that the effectiveness of retractions of misinformation was not affected by attitudes. That is, people's attitudes did not affect the extent

to which a retraction reduced their reliance on a piece of attitude-relevant misinformation. To reconcile this finding with existing literature, we suggest that the effectiveness of attitude-incongruent retractions will depend on whether or not accepting the retraction will induce a requirement to change the underlying attitude: When accepting a retraction does not require change in underlying attitudes, it will not be rejected for attitudinal reasons; when a retraction does challenge people's underlying attitudes, they will resist it.

References

- Aldy, J. E., Kotchen, M. J., & Leiserowitz, A. A. (2012). Willingness to pay and political support for a US national clean energy standard. *Nature Climate Change*, *2*, 596-599.
- Ayers, M. S., & Reder, L. M. (1998). A theoretical review of the misinformation effect: Predictions from an activation-based memory model. *Psychonomic Bulletin & Review*, *5*, 1-21.
- Batson, C. D. (1975). Rational processing or rationalization? Effect of disconfirming information on a stated religious belief. *Journal of Personality and Social Psychology*, *32*, 176-184.
- Berinsky, A. J. (2012). *Rumors, truths, and reality: A study of political misinformation*. Unpublished manuscript.
- Bizer, G. Y., Tormala, Z. L., Rucker, D. D., Petty, R. E. (2006). Memory-based versus on-line processing: Implications for attitude strength. *Journal of Experimental Social Psychology*, *42*, 646-653.
- Bjork, E.L., & Bjork, R.A. (1996). Continuing influences of to-be-forgotten information. *Consciousness and Cognition*, *5*, 176-196.
- Casiday, R., Cresswell, T., Wilson, D., & Panter-Brick, C. (2006). A survey of UK parental attitudes to the MMR vaccine and trust in medical authority. *Vaccine*, *24*, 177-184.
- Dechêne, A., Stahl, C., Hansen, J., & Wänke, M. (2010). The truth about the truth: A meta-analytic review of the truth effect. *Personality and Social Psychology Review*, *14*, 238-257.
- Devine P., & Elliot, A. (1995). Are racial stereotypes *really* fading? The Princeton Trilogy revisited. *Personality and Social Psychology Bulletin*, *21*, 1139-1150.

- Ecker, U. K. H., Lewandowsky, S., & Apai, J. (2011a). Terrorists brought down the plane! — No, actually it was a technical fault: Processing corrections of emotive information. *Quarterly Journal of Experimental Psychology*, *64*, 283-310.
- Ecker, U. K. H., Lewandowsky, S., Oberauer, K., & Chee, A. E. H. (2010b). The components of working memory updating: An experimental decomposition and individual differences. *Journal of Experimental Psychology: Learning Memory and Cognition*, *36*, 170-189.
- Ecker, U. K. H., Lewandowsky, S., Swire, B., & Chang, D. (2011b). Correcting false information in memory: Manipulating the strength of misinformation encoding and its retraction. *Psychonomic Bulletin & Review*, *18*, 570-578.
- Ecker, U. K. H., Lewandowsky, S., & Tang, D. T. W. (2010a). Explicit warnings reduce but do not eliminate the continued influence of misinformation. *Memory & Cognition*, *38*, 1087-1100.
- Ecker, U. K. H., Swire, B., & Lewandowsky, S. (in press). Correcting misinformation—A challenge for education and cognitive science. In D. N. Rapp & J. Braasch. *Processing Inaccurate Information: Theoretical and Applied Perspectives from Cognitive Science and the Educational Sciences*. Cambridge, MA: MIT Press.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191.
- Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, *54*, 297-327.

- Fielding, K. S., Head, B. W., Laffan, W., Western, M., & Hoegh-Guldberg, O. (2012). Australian politicians' beliefs about climate change: Political partisanship and political ideology. *Environmental Politics, 21*, 712-733.
- Gal, D., & Rucker, D. D. (2010). When in doubt, shout! Paradoxical influences of doubt on proselytizing. *Psychological Science, 21*, 1701-1707.
- Garrett, R. K., Nisbet, E. C., & Lynch, E. K. (in press). Undermining the corrective effects of media-based political fact checking? The role of contextual cues and naïve theory. *Journal of Communication*.
- Gilbert, D. T., Krull, D., & Malone, P. (1990). Unbelieving the unbelievable: Some problems in the rejection of false information. *Journal of Personality and Social Psychology, 59*, 601-613.
- Hart, P. S., & Nisbet, E. C. (2012). Boomerang effects in science communication: How motivated reasoning and identity cues amplify opinion polarization about climate mitigation policies. *Communication Research, 39*, 701-723.
- Hastie, R., Park, B. (1986). The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychological Review, 93*, 258-268.
- Jacks, J. Z., & Cameron, K. A. (2003). Strategies for resisting persuasion. *Basic and Applied Social Psychology, 25*, 145-161.
- Jacobson, G. C. (2010). Perception, memory, and partisan polarization on the Iraq War. *Political Science Quarterly, 125*, 31-56.
- Johnson, T. J., Bichard, S. L., & Zhang, W. (2009). Communication communities or “cyberghettos?”: A path analysis model examining factors that explain selective exposure to blogs. *Journal of Computer-Mediated Communication, 15*, 60-82.

- 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, D. S. (1993). Source monitoring. *Psychological Bulletin*, *114*, 3-28.
- Kendeou, P., & O'Brien, E. J. (in press). The Knowledge Revision Components (KReC) framework: Processes and mechanisms. In D. N. Rapp & J. Braasch (Eds.) *Processing Inaccurate Information: Theoretical and Applied Perspectives from Cognitive Science and the Educational Sciences*. Cambridge, MA: MIT Press.
- Kreitmann, N. (2006). Fantasy, fiction, and feelings. *Metaphilosophy*, *37*, 605-622.
- Kull, S., Ramsay, C., & Lewis, E. (2003). Misperceptions, the media, and the Iraq war. *Political Science Quarterly*, *118*, 569-598.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, *108*, 480-498.
- Kunda, Z., & Oleson, K. C. (1995). Maintaining stereotypes in the face of disconfirmation: Constructing grounds for subtyping deviants. *Journal of Personality and Social Psychology*, *68*, 565-580.
- Lamont, P. (2007). Paranormal belief and the avowal of prior scepticism. *Theory and Psychology*, *17*, 681-696.
- Lewandowsky, S., Ecker, U. K. H., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: Continued influence and successful debiasing. *Psychological Science in the Public Interest*, *13*, 106-131.
- 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.

- Loftus, E. F. (2005). Planting misinformation in the human mind: A 30-year investigation of the malleability of memory. *Learning & Memory, 12*, 361–366.
- Loken, B., & Hoverstad, R. (1985). Relationships between information recall and subsequent attitudes: Some exploratory findings. *Journal of Consumer Research, 12*, 155-168.
- Lord, C., Ross, L., & Lepper, M. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology, 37*, 2098-2109.
- Nyhan, B., & Reifler, J. (2010). When corrections fail: The persistence of political misperceptions. *Political Behavior, 32*, 303-330.
- Nyhan, B., & Reifler, J. (2011). *Opening the political mind? The effects of self-affirmation and graphical information on factual misperceptions*. Unpublished manuscript.
- Nyhan, B., Reifler, J., & Ubel, P. (2013). The hazards of correcting myths about health care reform. *Medical Care*. doi: 10.1097/MLR.0b013e318279486b
- Oberauer, K., & Vockenberg, K. (2009). Updating of working memory: Lingering bindings. *Quarterly Journal of Experimental Psychology, 62*, 967-987.
- Pedersen, A., Attwell, J., & Heveli, D. (2005). Prediction of negative attitudes toward Australian asylum seekers: False beliefs, nationalism, and self-esteem. *Australian Journal of Psychology, 57*, 148-160.
- Pedersen, A., Beven, J. P., Walker, I., & Griffiths, B. (2004). Attitudes toward Indigenous Australians: the role of empathy and guilt. *Journal of Community & Applied Social Psychology, 14*, 233-249.

Prasad, M., Perrin, A. J., Bezila, K., Hoffman, S. G., Kindleberger, K., Manturuk, K., . . .

Powers, A. S. (2009). "There must be a reason": Osama, Saddam, and inferred justification. *Sociological Inquiry*, *79*, 142-162.

Ratzan, S. C. (2010). Editorial: Setting the record straight: Vaccines, autism, and The Lancet. *Journal of Health Communication*, *15*, 237-239.

Redlawsk, D. P. (2002). Hot cognition or cool consideration? Testing the effects of motivated reasoning on political decision making. *Journal of Politics*, *64*, 1021-1044.

Redlawsk, D. P., Civettini, A. J. W., & Emmerson, K. M. (2010). The affective tipping point: Do motivated reasoners ever "get it"? *Political Psychology*, *31*, 563-593.

Richards, Z., & Hewstone, M. (2001). Subtyping and subgrouping: Processes for the prevention and promotion of stereotype change. *Personality and Social Psychology Review*, *5*, 52-73.

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.

Travis, S. (2010). *CNN poll: Quarter doubt Obama was born in U.S.* Retrieved from <http://politicalticker.blogs.cnn.com/2010/08/04/cnn-poll-quarter-doubt-president-was-born-in-u-s/>

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.

Wimshurst, K., Marchetti, E., & Allard, T. (2004). Attitudes of criminal justice students to Australian indigenous people: Does higher education influence student perceptions? *Journal of Criminal Justice Education, 15*, 327-350.

Young, G. (2010). Virtually real emotions and the paradox of fiction: Implications for the use of virtual environments in psychological research. *Philosophical Psychology, 23*, 1-21.

Appendix (to be placed in an online supplement)

Scenario of Experiment 1

‘Man Injured In Store Robbery’

Message 1. Police were called to investigate claims of an armed robbery on a liquor store late on Saturday night in the small town of Meera, approximately 30 minutes from Katherine in the Northern Territory.

Message 2. The call was made by a patron at the next door tavern who heard a disturbance and arrived to find the owner of the store on the floor behind the counter with an apparent head injury. The owner was the only person in the store.

Message 3. Police, who were coming from Katherine, arrived around 30 minutes after the call was made and secured the scene. An ambulance and paramedics arrived a few minutes later and gave first aid to the shop owner who was in severe shock, but in an otherwise stable condition.

Message 4. CCTV footage from a surveillance camera indicated that three men had invaded the store. They had approached the owner from behind and struck him over the head with a crowbar. The attack had lasted less than a minute in total and the footage indicated the attackers escaped with cash from the register, several bottles of spirits and cigarettes.

Message 5. (No-Misinformation): After reviewing the evidence from the CCTV video surveillance, police released a statement that they believed the three suspects were Caucasian youths in their late teens.

Message 5. (Misinformation): After reviewing the evidence from the CCTV video surveillance, police released a statement that they believed the three suspects were Aboriginal youths in their late teens.

Message 6. Based on the footage, the three suspects were described as male, around 19-20 years of age, of average height and slim build, who looked similar in appearance. They were all clothed in ragged shirts and shorts, but were not wearing any shoes.

Message 7. The forensic team and the major crime squad were called to the liquor store in question but all they found was an empty brown paper bag with paint stained finger prints. However, the finger prints were smudged and forensics could not get any useful evidence.

Message 8. Police interviewed some local residents who were at the tavern next to the liquor store but they claimed they would not be able to identify any of the attackers. They did report seeing a dark blue Holden Commodore speeding from the liquor store.

Message 9. Police found an abandoned car matching the description that had crashed into an embankment near the close-by community of Kooronga. Residents of the community were unhelpful in assisting with the investigation.

Message 10. Two days later, the same newspaper published the following.

Message 11 (No misinformation). Police released a second statement confirming that the owner of the store was the sole person in the store and no-one else was injured.

Message 11 (Retraction). Police released a second statement saying that after closer inspection of the CCTV footage, they no longer believed the suspects were of Aboriginal descent.

Message 11 (No retraction). Same as 'No misinformation.'

Message 12. The owner of the liquor store had been taken to hospital but was released the following morning. Police have since interviewed him, but as he was struck from behind he did not see the attackers. He did recall hearing a conversation between the attackers but he could not understand what they were saying.

Message 13. Local residents said that whilst this was a worrying incident, it was not uncommon and most liquor stores in the area had duress buttons that alerted police automatically to an incident.

Message 14. The police stressed that they would do everything in their power to apprehend the culprits, but that they feared the case would remain unresolved.

Open-ended Questionnaire of Experiment 1

Inference questions.

1. Why were all the attackers dressed in ragged clothes?
2. Why were none of the people from the tavern able to give a good description of the attackers?
3. Why did the shop owner have difficulty understanding the conversation between the attackers?
4. Why did the suspects appear similar to each other in the CCTV footage?
5. Why did the presumed getaway car crash into the embankment?
6. Why do most liquor stores in the area have a duress button?
7. Why were residents of the Kooronga community unhelpful in assisting the police?
8. Why would there have been paint stained finger prints on the evidence?
9. Why do the police fear the case will remain unsolved?

Fact-recall questions.

1. On what day did the incident occur?
2. Who made the call to police?
3. Which town were the police coming from?
4. How many attackers were there?

5. What weapon did the attackers use to hit the liquor store owner?
6. What did the attackers steal during the invasion?
7. What was the item that forensics found at the scene of the crime?
8. What sort of car was found abandoned?
9. How long did the liquor store owner stay in hospital for?
10. Who do you think the attackers were? (coded as inference question)

Retraction-awareness questions.

1. What was the purpose of the second police statement?
2. Was any of the information in the story subsequently corrected or altered? And if so, what was it?

Scenario of Experiment 2

‘Bank Robber Trapped in Storeroom’

Message 1. Police were lucky enough to apprehend the man responsible for an armed hold up in the eastern suburbs of Perth on Thursday morning. Officers say it may be the same man responsible for a string of robberies in neighbouring areas.

Message 2. Police arrived at the ANZ branch in Morley after being called to investigate claims of a hold up. When Police arrived they were startled to find the culprit locked up in the bank’s storeroom.

Message 3. A young employee, who had started at the bank only two weeks prior to the robbery, pressed the emergency button to alert police when she saw a man enter holding a gun.

Message 4. Police coming from the nearby Morley Police Station learnt that the robber had pulled out a gun demanding money, but the situation was under control before officers

arrived at the scene. Police released a statement the next morning detailing the heroic actions of one of the bank customers.

Message 5 (Misinformation). The customer, reportedly a local Aboriginal resident, had stepped in front of the main tellers and had convinced the robber to put down his gun. The manager of the bank managed to push the disarmed man into the storeroom until police arrived.

Message 6. There were three customers in the bank at the time of the robbery, two ladies and the young man. One of the ladies described how the robber waved the gun at the tellers and yelled for them to hand him money from behind the counter.

Message 7. Police and witnesses alike were shocked at the turn of events. Senior constable Matthew Finley said: “It was certainly an unexpected twist, we are very grateful to the gentleman who intervened, though we do not encourage such acts of bravery.”

Message 8. The manager of the bank told police that security screens, usually activated to protect tellers in such an event, were in the process of being fixed and that the attempted robbery came at the worst possible time.

Message 9. Witnesses told reporters from the Eastern Suburbs Reporter that the customer who prevented the robbery was uncomfortable and modest when giving his statement to police, which he chose to do at the scene instead of the police station.

Message 10. Although no one was injured during the incident, an ambulance was called to treat two of the bank clerks for shock. “We are very fortunate that no one was more seriously hurt,” said one paramedic.

Message 11 (Retraction). Police later released a second statement revealing that the man who helped apprehend the intruder was not an Aboriginal man as was first reported.

Message 11 (No retraction). Police later released a second statement confirming the intruder had been carrying a gun.

Message 12. Dianella resident Mary Sirolli (82), who was in the bank at the time, spoke to reporters: “I’m absolutely stunned” she said. “It’s something that you never expect to happen to you, I just can’t believe it.”

Message 13. The brave bank customer, who wished to remain anonymous, did tell reporters on the scene that he wasn’t acting brave to be a hero; he just didn’t think police would get there in time.

Message 14. Once all the witnesses had given police their statements, the assailant, who was reportedly a Caucasian man in his late thirties, was taken to the local police station where he will be questioned in regards to a number of robberies in the area.

Open-ended Questionnaire of Experiment 2

Inference questions.

1. Why did police say the course of events was ‘unexpected?’
2. Why was the young man uncomfortable about being interviewed after the event?
3. Why did the young man give his statement at the scene instead of the police station?
4. Why did the bank customer wish to remain anonymous?
5. What would be appropriate for the bank manager to do now?
6. What would be a better headline for the report of this incident?
7. What can people learn from this incident?
8. Why was the bank customer not scared of the robber?
9. How might the young man have managed to convince the robber to put down his gun?
10. Why did the elderly lady say she was ‘absolutely stunned’ when asked about the

incident?

Fact recall questions.

1. At which bank did the incident occur?
2. On which day did the incident occur?
3. Who alerted the police?
4. What police station were the police coming from?
5. How old was the robber said to be?
6. What did the bank manager say was being fixed at the bank?
7. How many customers were in the bank at the time of the robbery?
8. Why was the ambulance called?
9. What age was the elderly lady that told reporters that she was “absolutely stunned”?
10. What was the name of the Senior Constable that commented on the bravery of the young man?

Retraction-awareness questions.

1. What was the purpose of the second official police statement?
2. Was any of the information in the story subsequently corrected or altered? And if so, what was it?

Author Note

This research was facilitated by a Discovery Grant and an Australian Postdoctoral Fellowship to the first author, and a Discovery Grant and a Discovery Outstanding Researcher Award from the Australian Research Council to the second author. We thank Charles Hanich and Devon Spaapen for research assistance, and Nic Fay for suggesting the stereotype-incongruent scenario used in Experiment 2. The lab web address is <http://www.cogsciwa.com>.

Footnotes

¹ This is in contrast to another common usage of the term ‘misinformation’ in the literature on source memory and in particular eyewitness memory, where the term is used in a more general way to refer to erroneous information, and in particular post-event suggestive misinformation (cf. Loftus, 2005).

² For discussion of the reverse effects of memory on attitudes, see Bizer, Tormala, Rucker, and Petty (2006); Hastie and Park (1986); Loken and Hoverstad (1985).

³ Participants’ race or ethnicity was neither considered nor recorded in the participant selection process, and from publically available information it is estimated that about 80 % of participants were Caucasian and 20 % from culturally diverse (mainly Asian) backgrounds; only about 1 % could be expected to identify as Aboriginal.

⁴ The ATIA score of the high-prejudice group was on par with the population mean (2.85 on 0-6 scale) reported in Pedersen et al. (2004). The participants in Pedersen et al. came from the same city (Perth) but were on average much older (49.7 years) and less educated (less than half attending or having attended a tertiary institution) than the participants of the current study. Pedersen et al. reported a correlation between both age and education with racial prejudice, with younger and more educated people being on average less prejudiced. This means that our high-prejudice group cannot be described as extremely high in racial prejudice, but that the mean prejudice score is probably above-average for a student population.

⁵ Repeating these analyses excluding participants scoring below 2 on the fact-recall questions ($n = 3$) and participants from the retraction condition with a retraction-awareness score of 0 ($n = 6$) did not substantially alter the pattern of results.

⁶ In contrast to previous studies, where a central aspect of the scenario was retracted, such as the cause of a fire, the retraction in the present case concerned a relatively peripheral aspect of the scenario, and we hence expected a relatively low number of references to this critical piece of information.

⁷ For pragmatic reasons related to delays in ethics approval and project deadlines, pre-screening was done on three separate occasions, and participants were selected from the upper and lower quartiles of the three resulting distributions.

⁸ Repeating these analyses excluding participants scoring below 2 on the fact-recall questions ($n = 3$) and participants from the retraction condition with a retraction-awareness score of 0 ($n = 9$) did not substantially alter the result pattern.

Table 1

Racial prejudice (ATIA) scores, Experiments 1 and 2

Condition	Low-Prejudice Group		High-Prejudice Group	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Experiment 1				
No-misinformation	1.14	0.09	2.77	0.14
No-retraction	1.60	0.14	3.11	0.15
Retraction	1.25	0.11	2.85	0.12
Experiment 2				
No-retraction	1.45	0.12	4.18	0.23
Retraction	1.64	0.15	4.26	0.21

Note. ATIA, Attitudes Toward Indigenous Australians scale (Pedersen et al., 2004).

Table 2

Recall Accuracy, Experiments 1 and 2

Condition	Low-Prejudice Group		High-Prejudice Group	
Experiment 1	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
No-misinformation	0.62	0.04	0.63	0.03
No-retraction	0.69	0.05	0.74	0.03
Retraction	0.64	0.04	0.63	0.05
Experiment 2				
No-retraction	0.61	0.03	0.57	0.04
Retraction	0.63	0.02	0.56	0.04

Table 3

Contrasts, Experiment 1

Contrast		<i>F</i> (1,138)	<i>p</i>
Effects of Retraction			
1	<i>Low Racial Prejudice</i>	4.73	.03*
2	<i>High Racial Prejudice</i>	20.58	< .001*
Continued Influence Effects			
3	<i>Low Racial Prejudice</i>	12.11	.001*
4	<i>High Racial Prejudice</i>	9.27	.003*
Effects of Racial Prejudice			
5	<i>No-Misinformation</i>	< 1	
6	<i>No-Retraction</i>	8.53	.004*
7	<i>Retraction</i>	<1	

Note. * significant, Holm-Bonferroni corrected

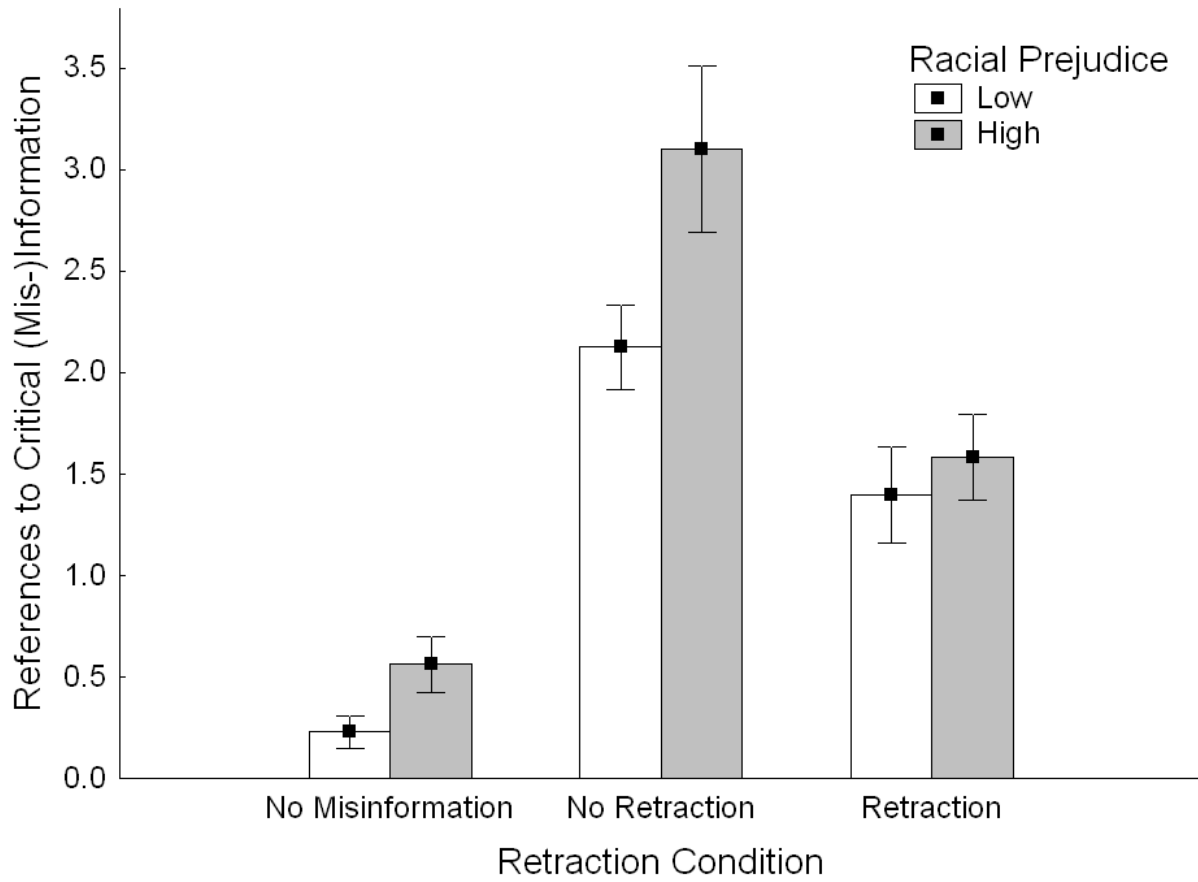


Figure 1. Mean number of references to the critical (mis-)information across conditions in Experiment 1. Error bars represent (unpooled) standard errors of the mean.

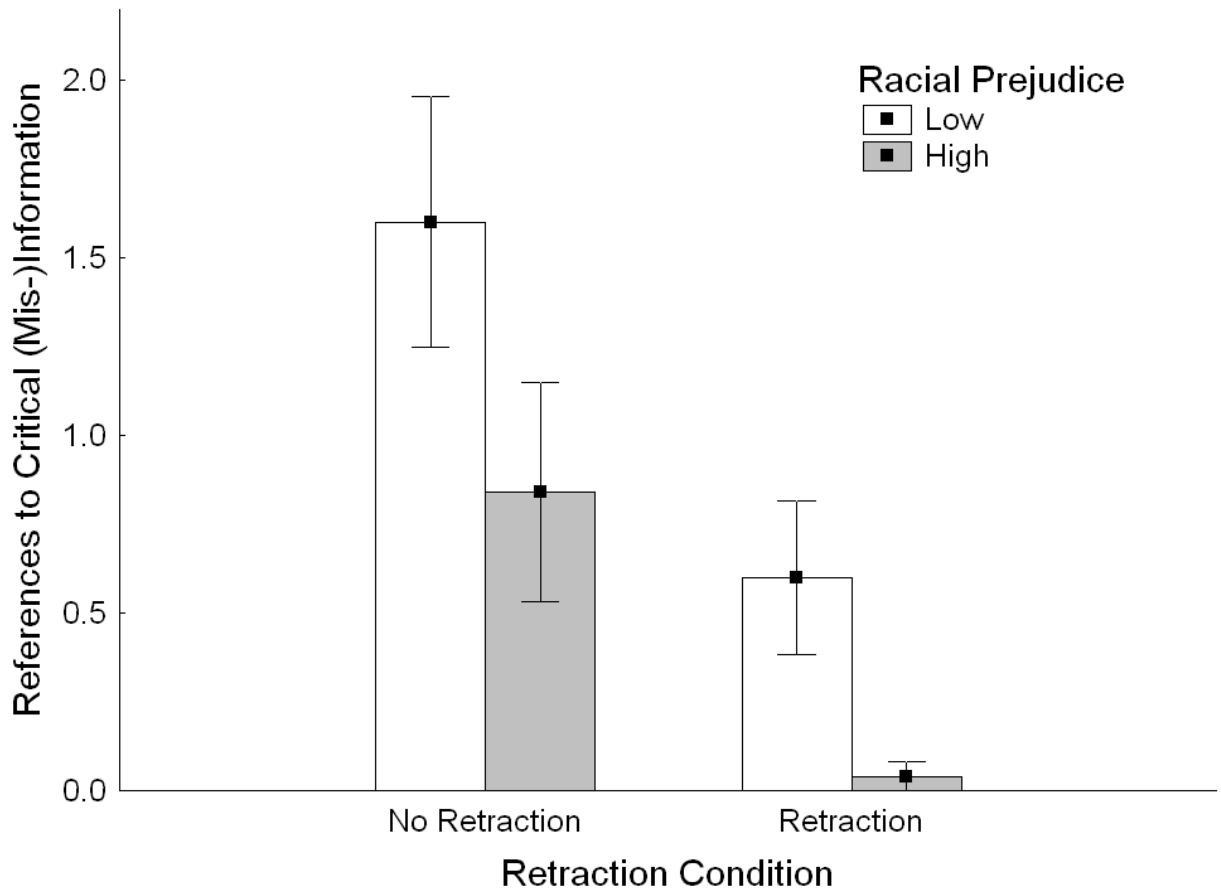


Figure 2. Mean number of references to the critical (mis-)information across conditions in Experiment 2. Error bars represent (unpooled) standard errors of the mean.