Published in final edited form as:

Psychol Sci. 2010 June ; 21(6): 801–803. doi:10.1177/0956797610371341.

Correcting False Memories

Lisa K. Fazio and Elizabeth J. Marsh

Duke University

While often impressive, memory is far from perfect. For example, the sentence “*The karate champion hit the cinder block*” is often misremembered as “*The karate champion broke the cinder block*” (Brewer, 1977). Hearing a list of related words such as “*bed, rest, tired ...*” leads people to claim “sleep” was presented when in fact it was not (Roediger & McDermott, 1995). Answering the question “*How fast was the white sports car going when it passed the barn while traveling along the country road?*” increases witnesses’ later reports of having seen a non-existent barn in an earlier video (Loftus, 1975, p. 566). These examples represent just a few of the many ways in which memory can go astray. Not only are these errors easily created, they often become vivid false memories held with high confidence. For example, false memories for non-presented words like “sleep” are so vivid that people often claim to remember which of two voices said the word (Payne, Elie, Blackwell, & Neuschatz, 1996).

False memories can be strikingly persistent. Warnings about memory errors are rarely effective (McDermott & Roediger, 1998), especially after the study phase (Greene, Flynn, & Loftus, 1982). Re-exposure to events is insufficient; hearing “*bed, rest, tired ...*” again reduces, but does not eliminate, false memories for “sleep” (McDermott, 1996; Watson, McDermott, & Balota, 2004). Even interventions that pinpoint specific contradictions between subjects’ memories and the original events are inadequate; many errors remain uncorrected even after subjects place an X next to each false memory (McConnell & Hunt, 2007).

Despite much evidence that false memories are difficult to correct, a finding from another literature allows a surprising prediction about the correction of false memories. In a prototypical experiment demonstrating the *hypercorrection effect*, participants answer general knowledge questions and rate their confidence in each response before viewing the correct answer. High-confidence errors are more likely to be corrected on a second test than are incorrect guesses (Butterfield & Metcalfe, 2001). Someone who strongly believes that Sydney is the capital of Australia will benefit more from the feedback “*Canberra*” than someone who simply guessed “*Sydney*.” The parallel in the false memory domain would be surprising, predicting that confidently-held false memories should be corrected more often than other errors. To examine this, we created false memories using sentences that encourage inferences; for example, “*The clumsy chemist had acid on his coat*” is often

misremembered as “*The clumsy chemist **spilled acid** on his coat.*” We examined ability to correct false memories as a function of initial confidence in the errors.

Method

Forty-six undergraduates studied 48 sentences (e.g., “*The karate champion hit the cinder block*”); each sentence was presented for 3500 ms and implied an action (e.g., *broke*). Materials were from McDermott and Chan (2006). On the initial test, subjects tried to complete each sentence fragment with the exact studied wording (e.g., “*The karate champion ____ the cinder block*”). On average, 2.08 words were needed to complete each sentence. Subjects rated their confidence using a 7-point scale, and then the original sentence re-appeared for 4 seconds with the previously missing portion bolded (e.g., “*The karate champion **hit** the cinder block*”). Subjects then participated in an unrelated experiment for 10 minutes, and finally were re-tested on all sentence fragments.

Results

Following McDermott and Chan (2006), each answer was categorized as correct (e.g., *hit*), a non-studied inference (e.g., *broke*), or another error (e.g., *kicked*). The on-line supplement details the scoring of responses, along with the distribution and calibration of confidence judgments on the initial test. Initially, inferences ($M = .51$) were more common than correct answers ($M = .25$), $t(45) = 8.23$, $SEM = .03$, or other errors ($M = .25$), $t(45) = 11.74$, $SEM = .02$. However, following feedback, participants corrected many of the errors. On the final test, participants completed more fragments correctly ($M = .74$) than with inferences ($M = .14$), $t(45) = 17.24$, $SEM = .03$ or other errors ($M = .12$), $t(45) = 19.89$, $SEM = .03$.

Our focus was on which errors were corrected on the final test, and whether there was hypercorrection of false memories. Figure 1 shows hypercorrection: high-confidence false memories were more likely to be corrected on the final test than were low-confidence memory errors. The within-subject gamma correlation between initial confidence and later correction was positive and significant, $\gamma = .13$, $t(44) = 2.05$, $SEM = .06$. The same pattern was obtained when inferences were combined with other errors, $\gamma = .14$, $t(45) = 2.51$, $SEM = .06$.

Discussion

With feedback, participants corrected more than two-thirds of their errors. Critically, we found hypercorrection of false memories: following feedback, subjects corrected more false memories (made with high confidence) than erroneous guesses. The implication for other episodic memory errors is that corrections will be most likely when feedback contradicts subjects' expectations.

Hypercorrection of false memories is consistent with the idea that people attend more to feedback when it is surprising (a hypothesis based on Rescorla and Wagner's (1972) animal learning model). Supporting this, memory for the feedback's color is improved following a high-confidence error (Fazio & Marsh, 2009; see also Butterfield & Metcalfe, 2006). An alternate explanation assumes the hypercorrection effect occurs because confidence in errors

is correlated with knowledge about the target domain. For example, most readers will be more confident when answering questions about Psychology than Chemistry. However, given that one makes an error about Psychology, the feedback will be better remembered because it can be associated to one's pre-existing knowledge. Returning to the present data, our paradigm controls for background knowledge. Although all false memories depend upon activation of meaning structures, this is uncorrelated with confidence in episodic memories. Our finding of hypercorrection for episodic memories means that differences in domain knowledge cannot be solely responsible for the hypercorrection effect.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Barbie Huelser and Aaron Johnson for help with data collection. This research was supported by the James S. McDonnell Foundation. The opinions expressed are those of the authors and do not represent the views of the foundation.

References

- Brewer WF. Memory for the pragmatic implications of sentences. *Memory & Cognition*. 1977; 5:673–678. [PubMed: 24203284]
- Butterfield B, Metcalfe J. Errors committed with high confidence are hypercorrected. *Journal of Experimental Psychology: Learning, Memory, & Cognition*. 2001; 27(6):1491–1494.
- Butterfield B, Metcalfe J. The correction of errors committed with high confidence. *Metacognition and Learning*. 2006; 1(1):69–84.
- Fazio LK, Marsh EJ. Surprising feedback improves later memory. *Psychonomic Bulletin & Review*. 2009; 16:88–92. [PubMed: 19145015]
- Greene E, Flynn MS, Loftus EF. Inducing resistance to misleading information. *Journal of Verbal Learning & Verbal Behavior*. 1982; 21:207–219.
- Loftus EF. Leading questions and the eyewitness report. *Cognitive Psychology*. 1975; 7:560–572.
- McConnell MD, Hunt RR. Can false memories be corrected by feedback in the DRM paradigm? *Memory & Cognition*. 2007; 35(5):999–1006. [PubMed: 17910183]
- McDermott KB. The persistence of false memories in list recall. *Journal of Memory & Language*. 1996; 35:212–230.
- McDermott KB, Roediger HL III. Attempting to avoid illusory memories: Robust false recognition of associates persists under conditions of explicit warnings and immediate testing. *Journal of Memory & Language*. 1998; 39:508–520.
- Payne DG, Elie CJ, Blackwell JM, Neuschatz JS. Memory illusions: Recalling, recognizing, and recollecting events that never occurred. *Journal of Memory & Language*. 1996; 35:261–285.
- Rescorla, RA.; Wagner, AR. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: Black, AH.; Prokasy, WF., editors. *Classical Conditioning II: Current Research and Theory*. New York: Appleton-Century-Crofts; 1972. p. 64–99.
- Roediger HL III, McDermott KB. Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition*. 1995; 21:803–814.
- Watson JM, McDermott KB, Balota DA. Attempting to avoid false memories in the Deese/Roediger-McDermott paradigm: Assessing the combined influence of practice and warnings in young and old adults. *Memory & Cognition*. 2004; 32(1):135–141. [PubMed: 15078050]

Figure 1.

Average proportion of the inferences on the initial test that were corrected on the final test, separated by initial confidence rating. Bolded line is the best fitting trend line.