Memory

ISSN: 0965-8211 (Print) 1464-0686 (Online) Journal homepage: https://www.tandfonline.com/loi/pmem20

Hypercorrection of high-confidence errors in the classroom

Shana K. Carpenter, Cynthia L. Haynes, Daniel Corral & Kam Leung Yeung

To cite this article: Shana K. Carpenter, Cynthia L. Haynes, Daniel Corral & Kam Leung Yeung (2018) Hypercorrection of high-confidence errors in the classroom, Memory, 26:10, 1379-1384, DOI: 10.1080/09658211.2018.1477164

To link to this article: https://doi.org/10.1080/09658211.2018.1477164

Published online: 19 May 2018.

MEMORY
2018, VOL. 26, NO. 10, 1379–1384
https://doi.org/10.1080/09658211.2018.1477164

REPORT

Hypercorrection of high-confidence errors in the classroom

Shana K. Carpenter[a], Cynthia L. Haynes[b], Daniel Corral[a] and Kam Leung Yeung[a]

[a]Department of Psychology, Iowa State University, Ames, IA, USA; [b]Department of Horticulture, Iowa State University, Ames, IA, USA

ABSTRACT

People often have erroneous knowledge about the world that is firmly entrenched in memory and endorsed with high confidence. Although strong errors in memory would seem difficult to "un-learn," evidence suggests that errors are more likely to be corrected through feedback when they are originally endorsed with high confidence compared to low confidence. This *hypercorrection effect* has been predominantly studied in laboratory settings with general knowledge (i.e., trivia) questions, however, and has not been systematically explored in authentic classroom contexts. In the current study, college students in an introductory horticulture class answered questions about the course content, rated their confidence in their answers, received feedback of the correct answers, and then later completed a posttest. Results revealed a significant hypercorrection effect, along with a tendency for students with higher prior knowledge of the material to express higher confidence in, and in turn more effective correction of, their error responses.

ARTICLE HISTORY
Received 12 January 2018
Accepted 11 May 2018

KEYWORDS
Hypercorrection; error correction; metacognition; memory; learning

Learning is a complex process that involves assimilation of new information with prior knowledge. What is already known about something can influence new learning, usually in a positive way such that a greater degree of prior knowledge predicts more effective learning of new information on that topic (Alexander, Kulikowich, & Schulze, 1994; Boscolo & Mason, 2003; McNamara, 2001).

But, what if the prior knowledge is false? It is not uncommon for people to have misconceptions in their knowledge, and sometimes these misconceptions can be quite strong. For example, Sadler, Schneps, and Woll (1987) demonstrated that a number of university graduates readily endorsed the false belief that seasonal changes on earth are caused by the distance between the earth and sun, rather than the tilt of the earth's axis. Other studies have reported this and other fairly common scientific misconceptions (e.g., see Atwood & Atwood, 1996). Such erroneous knowledge, particularly if it is highly accessible in memory, may *inhibit* – rather than facilitate – learning of new information.

Interestingly, however, errors endorsed with high confidence are more likely to be corrected than errors endorsed with low confidence – a finding called the *hypercorrection effect*. In the first study to demonstrate this effect, Butterfield and Metcalfe (2001) presented college students with general knowledge questions (e.g., What poison did Socrates take at his execution?), asked students to rate their confidence in their answer from 1 to 7, and then (if the answer was incorrect) immediately showed them the correct answer. On a posttest given a short time later, students answered the same questions again, and were more likely to remember the correct answers for questions they previously answered incorrectly with high confidence compared to low confidence.

The hypercorrection effect has been demonstrated many times in laboratory settings (e.g., Butler, Fazio, & Marsh, 2011; Butler, Karpicke, & Roediger, 2008, Experiment 1; Butterfield & Metcalfe, 2006; Fazio & Marsh, 2009; Metcalfe & Miele, 2014). The consistency of this rather counterintuitive finding suggests that strong errors can indeed be corrected when feedback of the correct answer is given. Such findings are of potential importance to education, as they could help inform instructional practices aimed at correcting students' errors about material they are learning.

Importantly, however, there is a lack of research exploring the hypercorrection effect in authentic educational contexts. Instead, studies of the hypercorrection effect are commonly based on college students learning trivia questions in laboratory settings. Recent work with educationally-relevant material – such as text passages over scientific phenomena – has not consistently demonstrated the hypercorrection effect (van Loon, Dunlosky, van Gog, van Merriënboer, & de Bruin, 2015). In the van Loon et al. study, the learning materials were based on information commonly taught in science curricula, raising the possibility that the hypercorrection effect may not occur for authentic course material.

On the other hand, there is reason to predict that hypercorrection of high-confidence errors does occur for course

CONTACT Shana K. Carpenter shacarp@iastate.edu Department of Psychology, Iowa State University, W112 Lagomarcino Hall, 901 Stange Road, Ames 50011–1041, IA, USA

material. Theoretical accounts of hypercorrection propose that the effect occurs due to a sense of surprise and heightened attention to feedback following a high-confidence error (Butterfield & Metcalfe, 2006; Fazio & Marsh, 2009), or the tendency for high-confidence errors to reflect temporarily inaccessible correct knowledge that was actually known at the time of the error – the "knew-it-all-along" effect (Metcalfe & Finn, 2011; Sitzman, Rhodes, & Tauber, 2014; Sitzman, Rhodes, Tauber, & Liceralde, 2015). Students taking a course over a given subject are likely to have a range of prior knowledge over that subject, and this prior knowledge could contribute to the correction of high-confidence errors. In particular, students with a higher degree of prior knowledge may be more confident in their wrong responses and thus feel a greater sense of surprise when they learn that they have made an error, or they may be more likely to have had some knowledge of the correct answer all along. Thus, students may readily show a hypercorrection effect for course knowledge, particularly if they are more knowledgeable about the material. Such possibilities remain unknown, as previous research has not explored the hypercorrection effect in actual course settings with course material.

The purpose of the current study was to explore students' errors in course knowledge, and their likelihood of correcting these errors as a function of confidence, in an actual course. Students in an introductory horticulture class answered questions about the course material, rated their confidence in their answers, and then received feedback of the correct answers. Students then completed a posttest over the same questions. A greater likelihood of providing correct answers on the posttest following high-confidence errors versus low-confidence errors would demonstrate the hypercorrection effect. To explore the potential role of prior knowledge, we examined the relationship between students' scores on the pretest and their likelihood of correcting pretest errors as a function of confidence.

Method

Students and course

Participants were 88 students enrolled in an introductory horticulture course at Iowa State University. The course covered topics related to growing plants in and around the home, maintaining gardens and landscapes, plant propagation, and caring for lawns, trees, shrubs, and flowers.

Materials

A set of 53 questions covering course concepts was prepared by the instructor. Twenty-six questions required a short-answer response (e.g., What causes blossom-end rot on tomatoes?), and 27 required a true/false response (e.g., You should add a layer of gravel to the bottom of a container to increase drainage). The questions were designed to address students' knowledge, including potential errors in their knowledge, associated with the course material.

Design and procedure

On the first day of class, students were invited by the instructor to complete a survey about their knowledge of course concepts. The survey was made available to students via the online course management system immediately after the first class, and students were given one week to complete it. During this one-week interval, students did not receive instruction over any of the concepts addressed in the 53 questions, and did not have access to the answers in any of their course materials.

The survey consisted of a pretest – during which students provided an answer to each of the 53 questions, rated confidence in their answer, and then received feedback – and a posttest during which students were shown the same questions again and asked to provide the correct answers, this time without receiving feedback. When students opened the link to the survey, they first saw instructions informing them that they would be asked some questions about their knowledge of concepts from the course. The instructions informed students that the survey responses would not be shared, and the responses were not graded. Students received course credit for completing the survey, regardless of their performance. The instructions also informed students that they should complete the survey individually, without help from books, notes, friends, or other sources, as students' individual responses would help provide diagnostic information that would assist in course planning.

Each of the 53 questions was presented one at a time, in a different random order for each student. Short-answer questions were displayed with a response box below the question where students entered their answers. True/false questions were displayed with the options "True" and "False" below the question, and a radio button next to each option. Students responded to questions at their own pace, and were required to enter a response to each question. If students left a question blank and tried to continue, an error appeared reminding them to provide an answer to the question before continuing.

Upon entering a response to the question, students saw a confidence rating scale appear below the question, inquiring "How confident are you in your answer?" with the response options 1 ("I am sure I am wrong") to 7 ("I am sure I am right"). After making a confidence judgment, the question and the correct answer were shown together, and students clicked a button marked "continue" to move on to the next question. For true/false questions correct answer feedback consisted of the word "true" or "false" presented below the question. For short answer questions (e.g., How can blossom-end rot be prevented?) feedback consisted of the brief response that provided a short answer to the question (e.g., water more frequently). All 53 questions were presented in this fashion – the question itself, the student's response, the student's confidence judgment, and then feedback.

After finishing the list of 53 questions, instructions on the screen informed students that they would now be tested over the correct answers to those same questions. Students were asked to do their best and to enter a response even if they were uncertain, and to again answer the questions individually without any help from external sources. On this posttest, students saw each of the 53 questions again, in a new random order, and responded at their own pace. This time they did not rate confidence, and feedback was not provided. Altogether, the survey took approximately 20 min to complete.

Results

Scoring

A total of 69 students completed the survey by the deadline, reflecting a response rate of 78%. The sample comprised 48% females and 52% males. Initial inspection of the data revealed that three students did not provide answers to the posttest questions. Data from these students were discarded, leaving 66 students in the sample.

Responses to the short-answer questions were hand-scored. A response was given a score of 0 if it was completely incorrect, and a score of 1 if it matched the expected correct answer provided by the instructor. Students' responses were inspected in blind fashion by the instructor, who awarded a score of 1 to answers that were correct but worded differently from the expected answer, and a score of .5 to answers that contained some of the correct information but were not completely correct. For example, an answer to the question "How can blossom-end rot be prevented?" received full credit if it matched the expected answer ("water more frequently") or was simply worded differently ("water more," "add more water"). Partially correct responses included "proper watering," and "watch how much you water," and incorrect responses included "fertilize more" and "warm weather."

Pretest to posttest improvement

Students' performance improved significantly from pretest to posttest for short-answer questions (respectively, $M = .18$, $SD = .10$; $M = .61$, $SD = .17$), $t(65) = 24.60$, $p < .001$, $d = 3.03$, and for true/false questions (respectively, $M = .49$, $SD = .10$; $M = .79$, $SD = .12$), $t(65) = 16.18$, $p < .001$, $d = 1.99$. For short answer questions that students answered correctly or partially correctly on the pretest, the proportion that they answered correctly or partially correctly again on the posttest was .93 (SD = .10). For short answer questions initially answered incorrectly, the proportion that students later answered correctly or partially correctly was .58 (SD = .19). For true/false questions, the proportion answered correctly on the pretest that were answered correctly again on the posttest was .85 (SD = .12), and the proportion answered incorrectly on the pretest that were answered correctly on the posttest was .73 (SD = .18).

Relation of confidence to accuracy on the pretest

Table 1 shows the proportion of questions answered correctly on the pretest as a function of confidence. For short answer questions, answers on the pretest that were given higher confidence ratings were also more likely to be correct. This relationship was not as apparent for true/false questions, likely due to the higher probability of guessing the answer correctly.

Error correction as a function of confidence

Table 2 shows the proportion of responses that were initially incorrect on the pretest and later corrected on the posttest as a function of pretest confidence. Proportions include both correct and partially correct responses on the posttest. On short answer pretest questions, students generally complied with the instructions to enter a response. However, on a small number of trials (less than 10% of responses across participants, on average), responses were entered that represented errors of omission (e.g., "I don't know," or "?"). On an individual basis, these items were removed from analyses, ensuring that the analyses examining error correction were based on errors of commission.[1]

Both types of questions revealed a *hypercorrection effect* – i.e., errors initially committed with higher confidence were more likely to be corrected than errors initially committed with lower confidence. Goodman-Kruskal gamma correlations (Goodman & Kruskal, 1954) computed for each student revealed a significant hypercorrection effect for both short-answer questions ($G = .11$, $SEM = .05$), $t(65)$

Table 1. Proportion of responses correct on the pretest as a function of confidence.

	Confidence rating						
	1	2	3	4	5	6	7
Short answer	.09 (.02)	.11 (.02)	.16 (.03)	.25 (.03)	.32 (.04)	.50 (.06)	.63 (.08)
True/False	.52 (.06)	.53 (.04)	.49 (.03)	.44 (.03)	.48 (.04)	.54 (.05)	.51 (.06)

Note: Proportions are not directly comparable across different ratings of confidence, as not all participants used the full range of the scale. The number of participants contributing to the proportions for short answer questions were 60, 61, 62, 60, 49, 44, and 28 for confidence ratings 1–7, respectively. The number of participants contributing to the same proportions for true/false questions were 32, 53, 59, 60, 61, 45, and 39, respectively. Standard errors are in parentheses.

Table 2. Proportion of initial errors corrected as a function of confidence.

	Confidence rating						
	1	2	3	4	5	6	7
Short answer	.54 (.04)	.55 (.04)	.54 (.04)	.58 (.04)	.72 (.05)	.80 (.07)	.86 (.07)
True/False	.63 (.07)	.73 (.05)	.66 (.05)	.73 (.03)	.77 (.05)	.76 (.07)	.95 (.04)

Note: Proportions are not directly comparable across different ratings of confidence, as not all participants used the full range of the scale and fewer initial errors occurred at higher confidence ratings. The number of participants contributing to the proportions for short answer questions were 59, 61, 59, 53, 45, 27, and 11 for confidence ratings 1–7, respectively. The number of participants contributing to the same proportions for true/false questions were 23, 43, 55, 57, 54, 34, and 28, respectively. Standard errors are in parentheses.

= 2.31, $p = .024$, and true/false questions ($G = .24$, SEM = .06), $t(58) = 3.93$, $p < .001$.[2] Combining short-answer and true/false questions for each student, the overall gamma correlations based on the entire question set were also significant, ($G = .18$, SEM = .04), $t(65) = 4.27$, $p < .001$.[3]

As a supplemental analysis, for each student the errors made on the pretest were divided into high-confidence (i.e., receiving a rating of 5, 6, or 7) and low-confidence errors (i.e., receiving a rating of 1, 2, or 3), and the proportion of each that was later corrected on the posttest was computed. As shown in Figure 1, a greater proportion of high-confidence errors compared to low-confidence errors was corrected, and this was true for both short answer questions, $t(48) = 3.82$, $p < .001$, $d = .55$, and true/false questions, $t(55) = 3.83$, $p < .001$, $d = .51$.

Prior knowledge and error correction

To explore the role of prior knowledge in error correction, we examined the relationship between students' scores on the pretest and their correction of errors on the posttest. This relationship was positive and marginally significant, $r(64) = .20$, $p = .108$, indicating that higher prior knowledge predicted a higher rate of error correction. We calculated students' mean confidence in their incorrect responses on the pretest, and examined whether confidence in these errors mediated the relationship between pretest accuracy and error correction (see Figure 2). A bootstrapping analysis using the 5000 resampling method with bias-corrected confidence interval and point estimate (MacKinnon, Lockwood, & Williams, 2004; Preacher & Hayes, 2008) revealed a significant indirect effect, $B = .186$, 95% CI [.028, .486]. The direct effect of pretest accuracy on error correction was not significant.

Discussion

The current study revealed a significant hypercorrection effect for students' course-related knowledge of horticulture concepts. This finding is consistent with a number of laboratory studies (Butler et al., 2008; Butterfield & Metcalfe, 2001, 2006; Fazio & Marsh, 2009; Metcalfe & Finn, 2011; Metcalfe & Miele, 2014), and shows that students' hypercorrection of high-confidence errors occurs in authentic educational settings as well.

These results reveal new insights into the role of prior knowledge in error correction. Students who had more knowledge of the material also had higher confidence in their errors, which in turn predicted a higher rate of error correction. This is consistent with research showing that

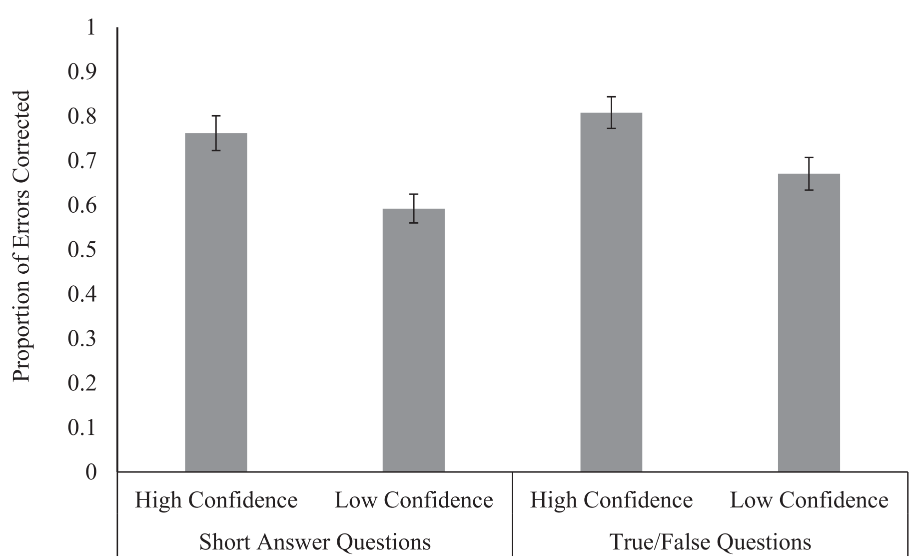

Figure 1. Proportion of high-confidence vs. low-confidence errors corrected for short answer questions (left) and true/false questions (right). Analyses include only students who made at least one high-confidence error (i.e., a rating of 5, 6, or 7) and one low-confidence error (i.e., a rating of 1, 2, or 3) for the short answer questions ($n = 49$) and for the true/false questions ($n = 56$). Error bars represent standard errors.

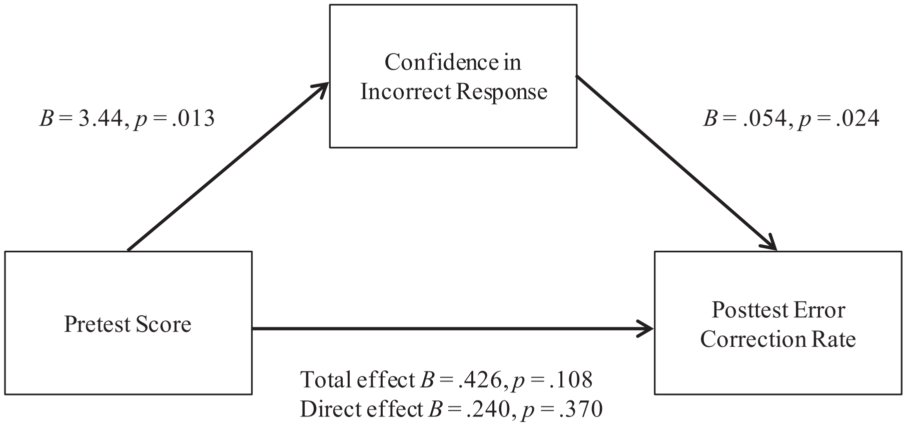

Figure 2. Mediation analysis of relationship between pretest score and posttest error correction through the mediator of confidence in pretest incorrect responses. All path coefficients are unstandardised.

people who are more knowledgeable about a given topic tend to be more overconfident in their knowledge on that topic (Son & Kornell, 2010). The positive association between prior knowledge and high-confidence error correction might also lend support to theories of the hypercorrection effect based on attention (Butterfield & Metcalfe, 2006; Fazio & Marsh, 2009) and the knew-it-all-along effect (Metcalfe & Finn, 2011; Sitzman et al., 2014, 2015), and suggests that prior knowledge could be leveraged to promote students' learning of correct information to replace highly intuitive but incorrect ideas.

Though the materials in the current study were more simple than some of the process-based misconceptions that have been documented (Sadler et al., 1987), the current results provide a novel demonstration of the hypercorrection effect with actual course knowledge that can guide future work on the correction of more complex scientific misconceptions, which is a natural and important extension of the hypercorrection effect (see Metcalfe, 2017). Some studies have shown mixed support for the hypercorrection effect with more complex beliefs and misconceptions (Fazio & Marsh, 2010; van Loon et al., 2015), indicating that complexity of the learning materials may be a key factor. Effective error correction may depend on the degree to which the error can be clearly noticed and refuted, which may be harder to do with complex materials. Indeed, van Loon et al. (2015) found that high-confidence errors to science questions followed by a reading passage were more likely to be corrected if the passage contained information that directly mentioned and refuted the error, as opposed to a passage that just contained the correct information without refuting the error.

In summary, students' high-confidence errors in course-related knowledge are effectively corrected through feedback. These results provide support for the idea that even strong errors can be replaced with correct information, and suggests that direct identification of these errors, followed by feedback, can be an effective way to learn course information. Future research should explore the extent to which the correction of errors occurs as a function of their strength and origin, and different types of course materials.

Notes

1. We also examined results when errors of omission were included. Gamma correlations between confidence in initial error responses (including both omission and commission errors) and later accuracy were significant for short answer questions ($G = .14$, SEM = .05), $t(65) = 2.99$, $p < .005$, and for the combined set of short answer and true/false questions ($G = .22$, SEM = .04), $t(65) = 4.45$, $p < .001$. Errors of omission were associated with very low confidence – so low, it appears, that students were not confident enough to enter a response – with 85% of these errors receiving a confidence rating of 1. For the students who made at least one omission error on the pretest ($n = 40$), the proportion of these errors corrected was .38, on average (SEM = .05). Thus, the hypercorrection effect was even stronger when including these very low confidence errors.
2. On the true/false questions, gamma could not be calculated for seven students due to invariant responses (i.e., all correct answers) on the posttest.
3. The same results emerged when excluding partially correct responses for short-answer questions (3.7% of responses) on the posttest. Including only fully correct and incorrect responses, gamma correlations were significant for the overall question set, $G = .19$ (SEM = .04), $t(65) = 4.13$, $p < .001$, and for short-answer questions, $G = .13$ (SEM = .06), $t(64) = 2.40$, $p = .019$.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This material is based upon work supported by the James S. McDonnell Foundation twenty-first Century Science Initiative in Understanding Human Cognition, Collaborative [grant number 220020483].

References

Alexander, P. A., Kulikowich, J. M., & Schulze, S. K. (1994). How subject-matter knowledge affects recall and interest. *American Educational Research Journal*, *31*, 313–337.

Atwood, R. K., & Atwood, V. A. (1996). Preservice elementary teachers' conceptions of the causes of seasons. *Journal of Research in Science Teaching*, *33*, 553–563.

Boscolo, P., & Mason, L. (2003). Topic knowledge, text coherence, and interest: How they interact in learning from instructional texts. *The Journal of Experimental Education*, *71*, 126–148.

Butler, A. C., Fazio, L. K., & Marsh, E. J. (2011). The hypercorrection effect persists over a week, but high-confidence errors return. *Psychonomic Bulletin & Review*, *18*, 1238–1244.

Butler, A. C., Karpicke, J. D., & Roediger, H. L. (2008). Correcting a metacognitive error: Feedback increases retention of low-confidence correct responses. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*, 918–928.

Butterfield, B., & Metcalfe, J. (2001). Errors committed with high confidence are hypercorrected. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 1491–1494.

Butterfield, B., & Metcalfe, J. (2006). The correction of errors committed with high confidence. *Metacognition and Learning*, *1*, 69–84.

Fazio, L. K., & Marsh, E. J. (2009). Surprising feedback improves later memory. *Psychonomic Bulletin & Review*, *16*, 88–92.

Fazio, L. K., & Marsh, E. J. (2010). Correcting false memories. *Psychological Science*, *21*, 801–803.

Goodman, L. A., & Kruskal, W. H. (1954). Measures of association for cross classifications. *Journal of the American Statistical Association*, *49*, 732–764.

MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research*, *39*, 99–128.

McNamara, D. S. (2001). Reading both high-coherence and low-coherence texts: Effects of text sequence and prior knowledge. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *55*, 51–62.

Metcalfe, J. (2017). Learning from errors. *Annual Review of Psychology*, *68*, 465–489.

Metcalfe, J., & Finn, B. (2011). People's hypercorrection of high-confidence errors: Did they know it all along? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*, 437–448.

Metcalfe, J., & Miele, D. B. (2014). Hypercorrection of high confidence errors: Prior testing both enhances delayed performance and blocks the return of the errors. *Journal of Applied Research in Memory and Cognition*, *3*, 189–197.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*(3), 879–891.

Sadler, P. M., Schneps, M. H., & Woll, S. (1987). *A private universe*. Santa Monica, CA: Pyramid Film and Video.

Sitzman, D. M., Rhodes, M. G., & Tauber, S. K. (2014). Prior knowledge is more predictive of error correction than subjective confidence. *Memory & Cognition*, *42*, 84–96.

Sitzman, D. M., Rhodes, M. G., Tauber, S. K., & Liceralde, V. R. (2015). The role of prior knowledge in error correction for younger and older adults. *Aging, Neuropsychology, and Cognition*, *22*, 502–516.

Son, L., & Kornell, N. (2010). The virtues of ignorance. *Behavioural Processes*, *83*, 207–212.

van Loon, M. H., Dunlosky, J., van Gog, T., van Merriënboer, J. J. G., & de Bruin, A. B. H. (2015). Refutations in science texts lead to hypercorrection of misconceptions held with high confidence. *Contemporary Educational Psychology*, *42*, 39–48.