The Role of Generating Versus Choosing an Error in Children's Later Error Correction

Abbey M. Loehr (Abbey.Loehr@Vanderbilt.edu)
Department of Psychology and Human Development, Vanderbilt University, 230 Appleton Pl #552
Nashville, TN 37203 USA

Lisa Fazio (Lisa.Fazio@Vanderbilt.edu)
Department of Psychology and Human Development, Vanderbilt University, 230 Appleton Pl #552
Nashville, TN 37203 USA

Bethany Rittle-Johnson (Bethany.Rittle-Johnson@Vanderbilt.edu)
Department of Psychology and Human Development, Vanderbilt University, 230 Appleton Pl #552
Nashville, TN 37203 USA

Abstract

Errors are common during learning, but what factors influence whether those errors are corrected? Evidence suggests error generation and memory for errors may be two important factors. Middle-school children studied and were tested on their memory for math definitions. After receiving correct answer feedback, children recalled their initial test answers before taking a final test. Memory for errors and error correction rates were higher for errors that were generated compared to errors that were chosen from a list. Further, memory for errors was positively correlated with error correction, even after controlling for age, grade, and math and reading skills. However, this relationship was only present for errors that were generated and not for errors that were chosen from a list. These findings suggest retrieval plays an important role in the relationship between memory for errors and error correction.

Keywords: error correction; retrieval practice; children; mathematics

Introduction

A growing body of research has demonstrated positive effects of committing errors during learning. For example, adults who received error management training when learning to use a computer program exhibited better transfer performance compared to adults who received error avoidant training (Keith & Frese, 2005). The error management training did not include any information about the task solution and encouraged adults to make errors and use feedback to learn from them. In contrast, the error avoidant training involved practice implementing step-by-step instructions. Relatedly, engaging students in exploratory problem-solving activities in which many errors are made relative to activities in which fewer errors are made supports retention and transfer (Bjork, 1994; Kapur & Bielaczyc, 2012).

Memory research has also shown a facilitative effect of generating errors so long as corrective feedback is given (see Metcalfe, 2017 for a review). The most direct evidence comes from studies comparing memory for the correct answer after participants are either instructed to generate an error or not for items in which the correct answer was highly unlikely to be guessed (e.g., slightly related word pairs). Participants exhibited better memory for the correct answer after generating an error before viewing the correct answer compared to studying the correct answer only (Kornell, Hays, & Bjork, 2009). Later studies replicated this effect with adults and elementary-school children (Carneiro, Lapa, & Finn, 2017) and identified important boundary conditions for the beneficial effects of generating errors. For example, the errors need to be at least somewhat semantically related to the correct answer (e.g., Huesler & Metcalfe, 2012; Kang et al., 2011), and the studied word pairs need to be related (Grimaldi & Karpicke, 2012; Knight et al., 2012). These boundary conditions suggest that active generation of errors may be helpful for learning by activating relevant knowledge, which may function as retrieval cues.

Role of Memory for Errors in Error Correction

While there is good evidence that committing errors can be beneficial for learning, less is known about what factors influence error correction. Most previous research has focused on adults' later memory for the correct answer after generating errors without examining memory for the previously committed errors. Thus, the potentially important role of memory for past errors in error correction remains relatively unknown.

Only two studies with adult participants have directly examined the relationship between accurate memory for previously committed errors (i.e., *memory for errors*) and error correction. Butler, Fazio, and Marsh (2011) asked adults to answer general knowledge questions and study correct-answer feedback. After a brief delay, adults answered the same general knowledge questions again. After this final test, the questions were re-presented and adults were asked to recall their initial test answers. Adults demonstrated very accurate memory for initially correct responses (98% recalled). Memory for errors was also good (85% of initial errors were accurately recalled). Importantly, adults corrected a greater proportion of errors if they accurately recalled the initial error than if they did not

accurately recall the error (.72 vs. .65). Following a similar procedure, Japanese adults in Iwaki, Nara, & Tanaka (2017) were tested on Kanji idioms, received correct-answer feedback, and were given a final test and asked to recall their initial test answers. Again, memory for errors facilitated error correction: a greater proportion of errors were corrected when adults remembered the error compared to when they did not (.80 vs. .54).

Generating and remembering errors may aid error correction for several reasons. First, retrieving related information (even if incorrect) within semantic networks creates a rich context for encoding the correct answer (Bjork, 1975; Kornell et al., 2009). Second, errors may act as retrieval cues if individuals associate the error with the correct answer (Barnes & Underwood, 1959; Carpenter, 2011; Pyc & Rawson, 2009). Relatedly, according to the theory of recursive reminding (Wahlheim & Jacoby, 2013), memory for errors may aid memory for the correct answer by acting as a reminder of the context in which an error was flagged as incorrect and the correct answer was presented. Finally, generating and remembering errors may increase the discrepancy between the error and corrective feedback. Individuals may pay more attention to this surprising feedback, which improves error correction (e.g., Fazio & Marsh, 2009).

Role of Age in Memory for Errors

While adults are good at recalling previous correct and incorrect responses (Finn & Metcalfe, 2008; Gardiner & Klee, 1976; Robinson & Kulp, 1970), evidence suggests children may exhibit worse memory for errors relative to previous correct responses. For example, 11-year-old children accurately identified .90 of initially correct responses and .63 of initial errors on a multiple-choice test after a one-week delay (Peeck et al., 1985). Thus, 11-year-olds demonstrated similar memory for errors after a one-week delay as adults in Butler et al. (2011). However, these rates should be considered an overestimation of children's memory for errors because they were asked to identify their previous errors from the multiple-choice options. It is unlikely children's recall rates will be as good.

Further, elementary-school children overestimate the accuracy of their past responses (Finn & Metcalfe, 2014). After studying social studies and science definitions and taking a cued recall test, 8- and 10-year-olds were shown the correct answer and asked to judge whether their answer on the initial test was correct or incorrect. Children misremembered 22% of their initial errors as having been correct. Thus, children's memory for errors may be faulty due to positively biased estimates of their past performance.

Current Study

Given the potentially important role memory for past errors is theorized to play in error correction, the current study examined the relationship between children's memory for their past errors and error correction. To measure children's memory for errors, children studied and were tested on math

definitions. After receiving correct-answer feedback, children recalled their initial test answers before taking a final test. Based on current evidence with adults, we predicted that better memory for errors would be related to greater error correction.

A second goal of the study was to determine if memory for errors and error correction were impacted by whether errors were generated or chosen from a list. Errors that were generated may have been endorsed with higher confidence than errors that were chosen from a list, which may reflect guesses. Adults were more likely to recall errors endorsed with higher confidence than errors endorsed with lower confidence (Butler et al., 2011). Further, research on the hypercorrection effect has demonstrated that high confidence errors are more likely to be corrected than low confidence errors (e.g., Butterfield & Metcalfe, 2001; Metcalfe & Finn, 2012). Therefore, if generated errors reflect errors endorsed with some confidence, children should demonstrate better memory for and correct more generated errors relative to errors chosen from a list.

The current study tests the generalizability of existing research and theory by examining children's memory for errors while studying math terms and their definitions – a new domain. Middle-school children (11- and 12-year-olds) were chosen for two main reasons. First, this age group is very similar to those included in both Finn & Metcalfe (2014) and Peeck et al. (1985), the two prior studies that examined children's memory for errors. Second, by age 12, children's performance on free recall tasks is quite similar to adults' (see Schneider & Pressley, 1997).

Method

Participants

Participants were 112 fifth-, sixth-, and seventh-grade children from three suburban private schools in Tennessee. Ten children were excluded from analysis either because their data were lost due to experiment software error, their session was interrupted, their responses indicated they did not take the task seriously, or they did not finish. The final sample included 102 children (M age = 11.9 yrs,; 66% female; 14% ethnic minorities). Children's scores on the previous year's standardized test were at the 76th percentile for math and 74th percentile for reading (scores for 7 children were not available). None of the children had severe reading disabilities and all children spoke English as their first language. Four children did not have enough time to complete the final recognition test.

Design and Procedure

The study occurred during a single one-on-one session lasting approximately 45 minutes and was conducted on a laptop computer. Children were told that they were going to learn definitions of math terms. They were told to study the definitions so that later they would be able to give the correct word for the definition on a test. Children wore headphones and heard all instructions and definitions read

aloud to control for any differences in reading ability between children.

During the first study-test cycle, children heard each term and its definition read aloud, which also appeared in writing on the screen. Each term was presented in red above its definition on the screen for 10 seconds. After the study phase, the definitions were shuffled and children took an immediate cued recall test. Children were encouraged to guess if they did not know an answer. If they could not guess they had the option of requesting a word bank to encourage guessing. When requested, the word bank appeared across the top of the screen and included a list of 36 math terms (see Figure 1). Twelve of these terms had not been presented during the study and test phases. The second study-test cycle immediately followed the first. Children re-studied the definitions and were asked to recall their initial test answer after the correct term disappeared. If children are unable to recall their initial answer, they were instructed to use the same word bank used during the testing phases to help them remember. No feedback was given during this phase. The definitions were shuffled again, and children took a final cued recall test as well as a final 4-option multiple-choice recognition test. The word bank was made available upon request during the final recall test to aid retrieval of the correct answer.

Materials

Children studied and were tested on 24 math definitions. The definitions were taken from the glossary of middle-school math textbooks so that few of the definitions would be known to children before the beginning of the study. See Figure 1 for an example of one of the definitions. The incorrect options for each definition on the final recognition test were randomly chosen from the list of terms included in the word bank.

Measured Outcomes

A response was scored as correct if it was spelled phonetically similar to the correct term. If a response was close to the correct term but was a different word (i.e., theory instead of theorem), it was scored as incorrect. Two primary outcome measures were calculated for each child. *Memory for errors* was calculated as the proportion of errors on the initial test that were accurately recalled. *Error correction* was calculated as the proportion of errors on the initial test that were corrected on the final cued recall and recognition tests.

Results

Word Bank Usage

Children often chose their answers from the word bank during both test phases and the recall past response phase. The word bank was requested for a greater proportion of trials on the initial and final tests than during the recall past response phase ($M = .64$, $SE = .02$ and $M = .55$, $SE = .03$ vs. $M = .39$, $SE = .02$, respectively). Additionally, a greater proportion of initial errors were chosen from the word bank than were generated ($M = .78$, $SE = .02$ vs. $M = .22$, $SE = .02$).

Initial and Final Test Performance

Children's performance increased from the initial test to both the final recall and recognition tests: initial cued recall test ($M = .35$, $SE = .02$), final cued recall test ($M = .46$, $SE = .02$), final recognition test ($M = .75$, $SE = .02$), $t(101) = 10.4$, $p < .001$, $t(97) = 30.9$, $p < .001$, respectively. Children corrected more of their initial errors on the recognition test ($M = .66$, $SE = .03$) than the cued recall test ($M = .28$, $SE = .02$).

arc	composite	heptagon	input	numeral	polyhedron	scalene	term	trapezoid
area	cone	histogram	matrix	parabola	prism	skew	tessellation	variable
chord	congruent	hypotenuse	mean	permutation	ray	slope	theorem	vertex
coefficient	conjecture	identity	net	plane	reciprocals	tangent	transformation	vertical

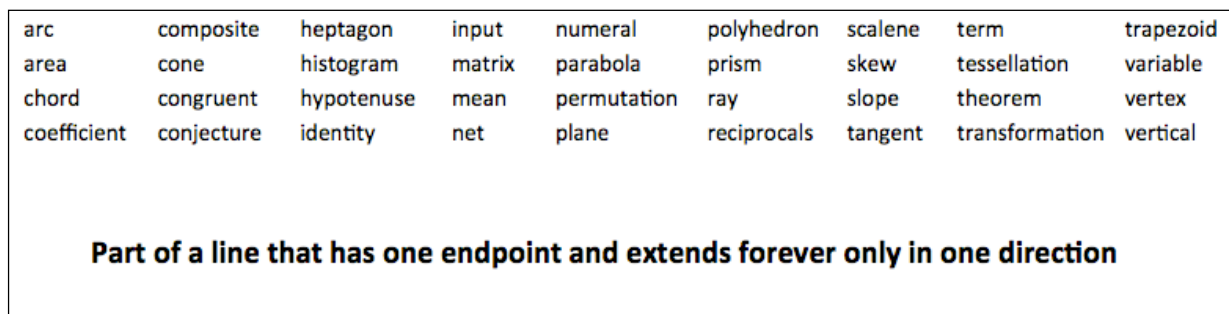**Part of a line that has one endpoint and extends forever only in one direction**

Figure 1. Example of definition for ray with word bank.

Initial and Final Tests For the initial and final cued recall tests, children saw each definition on the screen and heard it read aloud. They were asked to recall the correct term for each definition. Children typed their answers on the computer regardless of whether or not they requested the word bank. For the final recognition test, the incorrect options for each definition were randomly chosen from the list of terms included in the word bank.

Memory for Errors and Error Correction

While children accurately recalled many of their correct responses on the initial test ($M = .88$, $SE = .01$), memory for errors was much lower ($M = .28$, $SE = .02$). We explored whether memory for errors was low due to children's tendency to overestimate their past performance. Indeed, children inaccurately reported the correct answer instead of their past error for .20 ($SE = .02$) of their initial errors.

Similarly, Finn & Metcalfe (2014) found that 8- and 10-year-olds misremembered .22 of their initial errors as having been correct. Thus, 12-year-olds' memory for errors was positively biased to a similar extent.

Children's memory for errors was strongly, positively correlated with error correction on the final recall test, $r(100) = .46$, $p < .001$, and the final recognition test, $r(100) = .43$, $p < .001$. The correlations remained significant even after controlling for age, grade, and performance on the previous year's standardized math and reading tests, $r(89) = .26$, $p = .01$ and $r(89) = .25$, $p = .02$, respectively. Thus, children who were better at remembering their errors corrected more errors on the final tests, providing support for our hypothesis that memory for errors plays an important role in error correction.

Next, we examined memory for errors and error correction for errors that were generated versus errors that were chosen from the word bank. Given that both measured outcomes are proportions based on the total number of initial errors either generated or chosen, children who generated less than three errors and chose less than three errors from the word bank were excluded. This left 49 children who were included in the analysis[1]. Means and standard errors for memory for errors and error correction as a function of whether or not the errors were generated are presented in Table 1. As predicted, memory for errors was better for errors that were generated compared to errors that were chosen from the word bank, $t(48) = 5.1$, $p < .001$. Similarly, a greater proportion of errors that were generated were corrected on the final cued recall test than errors chosen from the word bank, $t(48) = 2.7$, $p = .01$. However, there was no difference in error correction on the final recognition test for errors that were generated versus chosen from the word bank, $t(48) = -.33$, $p = .75$.

Table 1: Memory for errors and error correction for generated and chosen errors.

	Generated M (SE)	Chosen M (SE)
Memory for Errors	.43 (.04)	.21 (.03)
Error Correction		
Cued recall	.32 (.04)	.23 (.02)
Recognition	.62 (.04)	.63 (.03)

Interestingly, there was a significant, positive correlation between memory for errors and error correction on the final cued recall and recognition tests for generated errors, $r(48) = .55$, $p < .001$, $r(48) = .35$, $p = .01$, respectively (see Figure 2). The correlation for the final cued recall test remained significant after controlling for age, grade, and performance on the previous year's standardized math and reading tests, $r(39) = .37$, $p = .02$, but not the recognition test, $r(39) = .08$, $p = .60$. In contrast, no relation between memory for errors and error correction on either the final cued recall or

___
[1] The same pattern of results was found when the full sample

was included in the analysis.
recognition test was found for errors that were chosen from the word bank, $r(47) = .07$, $p = .63$, $r(47) = -.01$, $p = .94$, respectively (see Figure 3).
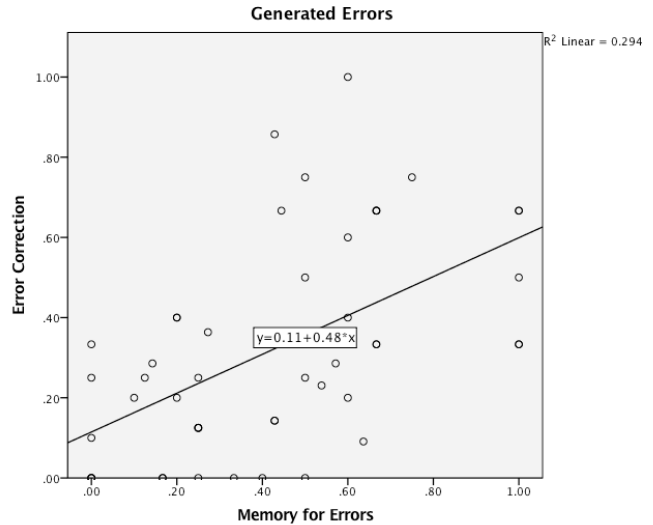

Figure 2: Scatterplot of memory for generated errors and error correction on the final recall test
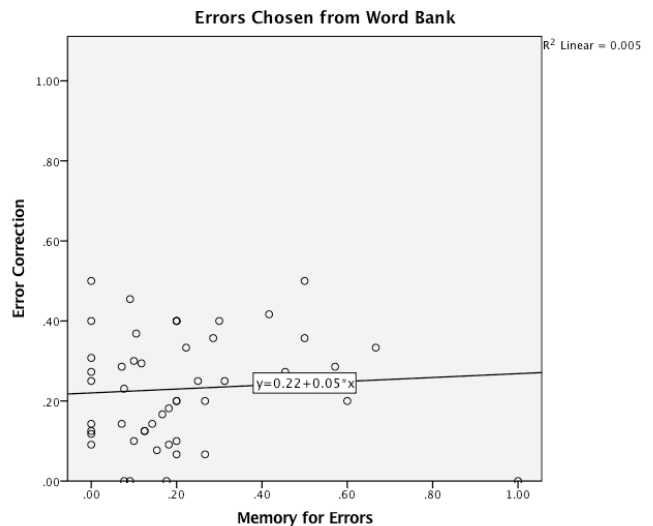

Figure 3: Scatterplot of memory for errors chosen from the word bank and error correction on the final recall test

Discussion

The current study is the first to examine children's memory for errors and its relation to error correction. After studying and being tested on above grade-level mathematics definitions, 12-year-old children reviewed the correct answers and recalled their previous test answers. Error correction was examined on a final recall and recognition

test. Children demonstrated good correct response memory, but their memory for errors was much lower than expected. In part, children's memory for errors was low because they misremembered 20% of their past errors as having been correct. Nearly as often as children accurately recalled their past errors, children inaccurately reported the correct answer instead of their past error. Similar to 8- and 10-year-olds in Finn & Metcalfe (2014), 12-year-olds in the current study seemed to exhibit a positivity bias and were overconfident in their past performance. These findings support research showing that children's metamemory monitoring, or the ability to self-monitor the accuracy of one's memories, continues to improve beyond middle childhood (Fandakova et al., 2017).

Despite poor memory for errors, children who were better at remembering their errors corrected more errors on both the final recall and recognition tests. In line with adult memory research (e.g., Butler et al., 2011; Iwaki et al., 2017), memory for errors seems to play an important role in error correction for middle-school children. Memory for errors may reflect an important metacognitive skill that guides attention to feedback and has the potential to impact future study behaviors. Indeed, Fandakova et al. (2017) found reciprocal relations between metamemory monitoring and cognitive ability, suggesting monitoring the accuracy of previous performance plays a critical role in learning.

Memory for errors, error correction, and the relation between the two varied depending on whether errors were generated or chosen from a list. Children were better at remembering errors that were generated and were also more likely to correct generated errors. Further, the positive relationship between memory for errors and error correction was only present for generated errors. Past work has shown a facilitative effect of error generation on error correction (e.g., Kornell et al., 2009), suggesting retrieval plays an important role in error correction. Indeed, results from the current study support this idea. The benefits of error generation influenced both memory for errors and its relation to error correction. Generating an error may make it more likely to be recalled later. In turn, retrieving past errors may increase the memory strength of the past error as well as support encoding of the correct answer by activating related information within semantic networks (Bjork, 1975; Kornell et al., 2009). Alternatively, according to reconsolidation theory (Nadar, Schafe & LeDoux, 2000; Lee, 2008), successfully retrieving an error provides an important opportunity for the error to be corrected.

Findings from the current study align with the hypercorrection effect (Butterfield & Metcalfe, 2001). Generated errors were more likely to be corrected than errors chosen from the word bank. A child's ability to generate an error may reflect that they held at least some level of confidence in that error. In contrast, errors chosen from the word bank may reflect lower confidence guesses. However, it is important to acknowledge that word bank usage may be an indicator of several other things besides confidence. For example, some children may simply request the word bank to check an answer, regardless of their confidence in the answer. Data collection is currently underway for a follow-up study directly measuring children's confidence in their initial test answers. This will allow us to disentangle word bank usage from confidence. Memory for errors may only aid error correction for errors in which that are endorsed with at least some confidence compared to errors that are wild guesses, regardless of whether the word bank was used.

In conclusion, findings from the current study provide support for memory for errors and error generation as important factors that influence error correction. Given the considerable variability in how teachers handle students' errors (Schleppenbach, Flevares, Sims, & Perry, 2007), research is needed to make recommendations about how to promote error correction. Findings from the current study suggest that committing errors is not detrimental to learning. In fact, providing opportunities for error generation, as well as attending to and remembering when errors were made, may aid learning.

However, more research is needed to determine how the nature of errors might interact with the effects of error generation and memory for errors. For example, research on conceptual change learning in science and mathematics domains warns against committing and practicing errors. Specifically, errors that reflect common misconceptions are persistent and difficult to change (e.g., Alvermann & Hynd, 1989; Vamvakoussi & Vosniadou, 2004). In fact, misconception errors often need to be directly disputed in order for learning to occur (e.g., Durkin & Rittle-Johnson, 2012). Relatedly, memory research on proactive interference has shown that well-learned information interferes with learning of related information (see Anderson & Neely, 1996). Thus, generation of and memory for misconception errors may hinder error correction.

Acknowledgments

References

Alvermann, D. E., & Hague, S. A. (1989). Comprehension of counterintuitive science text: effects of prior knowledge and text structure. *Journal of Educational Research, 82*, 197-202.

Anderson, M. C., & Neely, J. H. (1996). Interference and inhibition in memory retrieval. In E. L. Bjork & R. A. Bjork (Eds.), Handbook of perception and cognition: Vol. 10. Memory (2nd ed., pp. 237– 313). San Diego: Academic Press.

Barnes, J. M., & Underwood, B. J. (1959). "Fate" of first-list associations in transfer theory. *Journal of Experimental Psychology, 58*, 97–105.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). Cambridge, MA: MIT Press.

Bjork, R. A. (1975). Retrieval as a memory modifier: An interpretation of negative recency and related phenomena. In R. L. Solso (Ed.), Information processing and cognition: The Loyola Symposium (pp. 123–144). Hillsdale, NJ: Lawrence Erlbaum.

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*(6), 1238–1244.

Butterfield, B., & Metcalfe, J. (2001). Errors committed with high confidence are hypercorrected. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(6), 1491-1494.

Carneiro, P., Lapa, A., & Finn, B. (2018). The effect of unsuccessful retrieval on children's subsequent learning. *Journal of Experimental Child Psychology, 166,* 400-420.

Carpenter, S. (2011). Semantic information activated during retrieval contributes to later retention: Support for the mediator effectiveness hypothesis of the testing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(6), 1547-1552.

Durkin, K., & Rittle-Johnson, B. (2012). The effectiveness of using incorrect examples to support learning about decimal magnitude. *Learning and Instruction*, *22*(3), 206-214.

Fandakova, Y., et al. (2017). Changes in ventromedial prefrontal and insular cortex support the development of metamemory from childhood into adolescence. *Proceedings of the National Academy of Sciences, 114*(29), 7582-7587.

Fazio, L. K. & Marsh, E. J. (2009). Surprising feedback improves later memory. *Psychonomic Bulletin & Review, 16*(1), 88-92.

Finn, B., & Metcalfe, J. (2008). Judgments of learning are influenced by memory for past test. *Journal of Memory and Language, 58*(1), 19-34.

Finn, B., & Metcalfe, J. (2014). Overconfidence in children's multi-trial judgments of learning. *Learning and Instruction, 32*, 1-9.

Gardiner, J. M., & Klee, H. (1976). Memory for remembered events: An assessment of output monitoring in free recall. *Journal of Verbal Learning and Verbal Behavior, 15*(2), 227-233.

Grimaldi, P. J., & Karpicke, J. D. (2012). When and why do retrieval attempts enhance subsequent encoding? *Memory & Cognition, 40*(4), 505-513.

Huelser, B. J., & Metcalfe, J. (2012). Making related errors facilitates learning, but learners do not know it. *Memory & Cognition, 40*(4), 514-527.

Iwaki, N., Tomomi, N., Tanaka, S. (2017). Does delayed corrective feedback enhance acquisition of correct information? *Acta Psychologica, 181,* 75-81.

Kang, S. H., Pashler, H., Cepeda, N. J., Rohrer, D., Carpenter, S. K., & Mozer, M. C. (2011). Does incorrect guessing impair fact learning? *Journal of Educational Psychology, 103*(1), 48-59.

Kapur, M., & Bielaczyc, K. (2012). Designing for productive failure. *Journal of the Learning Sciences, 21*(1), 45–83.

Keith, N., & Frese, M. (2005). Self-regulation in error management training: Emotion control and metacognition as mediators of performance effects. *The Journal of Applied Psychology*, *90*(4), 677–91.

Knight, J. B., Ball, B. H., Brewer, G. A., DeWitt, M. R., & Marsh, R. L. (2012). Testing unsuccessfully: A specification of the underlying mechanisms supporting its influence on retention. *Journal of Memory and Language, 66*(4), 731-746.

Kornell, N., Hays, M. J., & Bjork, R. A. (2009). Unsuccessful retrieval attempts enhance subsequent learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(4), 989-998.

Lee, J. L. C. (2008). Memory reconsolidation mediates the strengthening of memories by additional learning. *Nature Neuroscience, 11*, 1264-1266.

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

Metcalfe, J., & Finn, B. (2012). Hypercorrection of high confidence errors in children. *Learning and Instruction, 22*(4), 253-261.

Nader, K., Schafe, G., & LeDoux, J. E. (2000). Fear memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature, 406*, 722–726.

Peeck, J., van den Bosch, A. B., & Kreupeling, W. J. (1985). Effects of informative feedback in relation to retention of initial responses. *Contemporary Educational Psychology, 10*(4), 303-313.

Pyc, M. A., & Rawson, K. A. (2009). Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *Journal of Memory and Language, 60*, 437–447.

Robinson, J. A., & Kulp, R. A. (1970). Knowledge of prior recall. *Journal of Verbal Learning and Verbal Behavior, 9*(1), 84-86.

Schleppenbach, M., Flevares, L. M., Sims, L. M., & Perry, M. (2007). Teachers' responses to student mistakes in Chinese and US mathematics classrooms. *The Elementary School Journal, 108*(2), 131–147.

Schneider, W., & Pressley, M. (1997). *Memory development between two and twenty.* Mahwah, NJ: Erlbaum.

Vamvakoussi, X., & Vosniadou, S. (2004). Understanding the structure of the set of rational numbers: a conceptual change approach. *Learning and Instruction, 14*(5), 453–467.

Wahlheim, C. N. & Jacoby, L. L. (2013). Remembering change: The critical role of recursive remindings in proactive memory. *Memory & Cognition,* 41(*1*), 1–15.