Retrieval-based Metacognitive Monitoring in Self-Regulated Learning

Mengjiao Wu (mwu8@kent.edu)
Department of Education Psychology, 150 Terrace Drive
Kent State University, Kent, Ohio 44224 USA
College of Foreign Languages,1550 Haigang Avenue
Shanghai Maritime University, Shanghai, 201306 China

Abstract

Metacognitive monitoring plays an important role in self-regulated learning. Accurate metacognitive monitoring facilitates effective control, which affects learning outcomes. Most studies that explore metacognitive monitoring have investigated learners' monitoring abilities when learners are explicitly cued to monitor. However, in real-world educational settings, learners are more commonly cued to control their learning. The primary goal of the current study was to investigate whether learners monitor their learning processes using retrieval when explicitly cued to control. Two experiments were conducted in pursuit of this goal. In the experiments, participants were instructed to learn Swahili-English word-pairs. Their learning performance was tested in subsequent cued-recall tests. Results suggest retrieval is likely practiced when learners are explicitly cued to control, but at a lower frequency or a more shallow level than when learners are explicitly cued to retrieve. In addition, the current study reported attempts to measure retrieval-based metacognitive monitoring using objective and online methods.

Keywords: metacognition; monitoring; control; retrieval; self-regulated learning;

Introduction

Metacognitive Monitoring

Metacognition refers to the cognition and control of one's own cognitive activities. It plays an essential role in self-regulated learning outcomes (Dunlosky & Metcalfe, 2009; Koriat, 2007). According to Dunlosky & Metcalf (2009), metacognition consists of three components: metacognitive knowledge, metacognitive monitoring, and metacognitive control. Metacognitive knowledge refers to people's declarative knowledge about cognition. Metacognitive monitoring involves the evaluation of an ongoing cognitive process (i.e., how well information has been learnt and the perceived likelihood of recalling the information in the future), whereas metacognitive control refers to the regulation of cognitive activities (Dunlosky & Metcalfe, 2009). Both the metacognition (Dunlosky & Metcalfe, 2009; Nelson & Narens, 1990; Thiede, Anderson, & Therriault, 2003) and self-regulated learning frameworks (Koriat, 2007; Nelson & Leonesio, 1988; Winne & Hadwin, 1998) postulate that monitoring informs control, which thereafter influences learning. It suggests effective monitoring is the prerequisite of effective control. The current study addresses metacognitive monitoring.

Research demonstrates that metacognitive monitoring can support learning outcomes (Zimmerman & Kitsantas, 1999; Maki, 1998; Dunlosky et al., 2005). Monitoring cues can support learning because they prompt the learner to monitor and evaluate their cognitive activities (e.g., make judgments of learning), whereas control cues can prompt learners to regulate their cognitive activities (e.g., make a study decision or change a study strategy). However, most studies on metacognitive monitoring were conducted using explicit metacognitive cues. That is, the metacognitive monitoring is performed under external instruction to do so for the sake of monitoring one's own learning (Dunlosky & Nelson, 1992; Nietfeld, Cao, & Osborne, 2005; Roebers, Krebs, & Roderer, 2014; Thiede, Anderson, & Therriault, 2003). It is unclear whether metacognitive monitoring is still performed and demonstrates the same impact on learning when explicit monitoring cues are not provided.

However, in many real-world educational settings, students are often presented with control cues rather than monitoring cues. For example, a language teacher may ask a student whether she wants to practice the new grammar point with more exercises. In this case, the student is explicitly cued to control (make a study decision) instead of monitoring. Do learners still monitor learning even when they are not cued to monitor but cued to control? Since effective control is dependent on accurate monitoring (Nelson & Narens, 1990; Renner & Renner, 2001), examining learners' metacognitive monitoring performance when there is no explicit cue for monitoring has important theoretical and practical implications for learning and education. That is, even if a learner has the ability to monitor his/her learning effectively when prompted, it doesn't necessarily imply that he/she will still engage in metacognitive monitoring when that monitoring cue is absent. Poor metacognitive control and low performance may owe itself more to the lack of engagement in metacognitive monitoring rather than to the abilities to monitor. As a result, it is important to investigate learners' metacognitive monitoring when no explicit monitoring cue is present.

This Study

The current study investigated college students' metacognitive monitoring in self-regulated learning when cued to control (i.e., asked to make study decisions). Judgments of learning (JOLs) are a measure of individuals' subjective evaluation of their own learning. JOLs made based on retrieval results (retrievability of the target) are found to be more accurate than the JOLs made based on other criteria (Metcalfe & Finn, 2008; Karpicke & Smith, 2012). In addition to that, retrieval practice is also a highlyinvestigated mnemonic strategy that contributes to better learning and memory performance (Landauer & Bjork, 1978; Rawson & Dunlosky, 2011). Thus, the focus of this study was narrowed down to the retrieval-based metacognitive monitoring. This study aimed to determine whether learners monitor their learning using the retrieval strategy when cued to control during a self-regulated learning paradigm, as well as its impact on learning and memory performance.

Existing assessments of metacognitive monitoring may have two methodological weaknesses: (1) the dependence on subjective self-report by participants (Nelson & Dunlosky, 1991); and (2) the use of offline measures (Bryce, Whitebread, & Szucs, 2014). Commonly used measures of metacognitive monitoring are self-reported in which participants are asked to make subjective judgments of their own cognitive processes (i.e., JOLs ask participants to rate how well they have learned the target materials). This could reduce the reliability of the measurement. Offline measures might not be able to reflect what occurs during the task because they are usually remote from the target cognitive task (i.e., delayed JOLs are made a period after the learning). Therefore, the secondary goal of this study was to report the attempts of measuring retrieval-based metacognitive monitoring using online and objective measures. It was hypothesized that learners will: (1) monitor their learning using retrieval strategies when cued to control in their self-regulated learning; and (2) the retrieval-based metacognitive monitoring of learning is measurable with online and objective measures.

Experiment 1

Method

Participants Participants included thirty-nine sophomore or junior level college students enrolled in an educational psychology course at a public university in Northeast Ohio. Students received course credits for participation.

Stimuli All stimulus presentations were programmed and administered on a computer using the custom program, E-prime 2.0. All participants were assigned to, and seated in front of a 15-inch desktop computer, with up to four participants present at the same time. The stimuli included the Swahili-English paired associates published by Nelson and Dunlosky (1994), which were normed by difficulty level. The Swahili-English word-pairs were used as learning and testing materials in this study (e.g., *dunia-world*).

Procedure At the beginning of the experiment, participants were provided a brief overview of the experiment, including being told they would study 70 Swahili-English word-pairs in this experiment. After two practice trials, participants were told that the main task would start. Their task goal was to learn as many word-pairs as possible and recall at least 70% of the word-pairs in the final test. The main task was consisted of a learning phase, a global judgement of learning (a global JOL), and testing. The learning phase was composed of 70 learning trials. In each trial, participants were presented with a Swahili-English word-pair for 8 seconds. Then they were instructed to solve a distractor task and make a study decision regarding the next study step. Each item was presented after a 500-millisecond delay, in which, a blank screen with a black "+" in the middle of the screen was presented. The presentation of a word-pair was immediately followed by a distractor task which was used to prevent participants from rehearsing the item. The distractor task was a two-digit mental addition task. Participants were instructed to type in the answer to the addition question and press "Enter" key on the keyboard to submit the answer within 20 seconds. Once an answer was submitted, the screen would indicate whether it was correct or incorrect for 1500 milliseconds, and then move to the study-decision intervention. A lack of response within 20 seconds would be treated as an incorrect response. In the self-paced study-decision intervention, participants were presented with two options placed horizontally in the middle of the computer screen, "Study Again" on the left side and "Next" on the right side. Participants were instructed to press the "F" key on the keyboard to see the same item for another 4 seconds, and press "J" key to advance to the next item. The study-decision response time (RT) was the time it took from the onset of the stimuli of the study-decision intervention, to the time a decision was submitted.

Upon completion of all the learning trials, participants were prompted to take a self-paced global judgment of learning (JOL) on their overall learning performance, rating how well they thought they had learned all the 70 items on a scale ranging from 0% to 100%. They were instructed to input a number from 1 to 100 to respond and press "Enter" key to submit the JOL response. In the end, a cued-recall test was used to assess participants' learning performance of the Swahili-English word-pairs. The pairs were presented in a random order. In testing, the Swahili word was presented on the left side of the screen (the cue; e.g., *dunia* –?). Participants needed to recall the English equivalent translation (e.g., *world*), type the answer into the designated area on the screen, and submit their responses by pressing the "Enter" key. A lack of response within 20 seconds would be treated as an incorrect recall. An item was scored as correct or incorrect by matching the typed response with the correct English target. The typed answers were also manually checked and rescored by researchers after the experiment. If the response and the correct answer matched morphologically or semantically, the response was counted as correct. After participants completed the final cued-recall test, they were thanked and dismissed. All participants completed the assessment within an hour.

Results

Due to the limit of words, only the most important results are reported in this document. In the final cued-recall test, every participant was tested on their memory of the 70 Swahili-English word-pairs. In the test trials, a correct response was coded as 1, and an incorrect response was coded as 0. Thus, the mean test accuracy for each participant ranged from 0 to 1, representing the proportion of correct recalls out of the total 70 test trials (M = .16, SD = .13). When a participant decided to restudy an item, the decision was re-coded as 0. When a participant decided to advance to the next item, the decision was re-coded as 1. Thus, the aggregated study-decision responses ranged from 0 to 1, representing the proportion of the decision of advancing to the next item without restudy. The means and standard deviations of the test accuracy, study-decision response, study-decision response time, and global JOL are presented in Table 1.

Table 1: Means and standard deviations of the test accuracy, study-decision response, study-decision response time, and global JOL in Experiment 1.

	n	*M*	*SD*
Test accuracy	39	.16	.13
Global JOL	38	25.34	15.02
Study-decision response	39	.79	.26
Study-decision RT (ms)	39	1118.82	596.69

Correlation analysis showed study-decision RT was positively correlated to test accuracy ($r(39) = .34$, $p = .034$), and global JOLs ($r(38) = .34$, $p = .04$); study-decision RT was negatively correlated to study-decision response ($r(39) = -.33$, $p = .04$). The study decision, on the other hand, was only correlated with study-decision RT, but had no correlation with test accuracy, or global JOL.

In order to examine whether study decision and study-decision RT could be used to explain the variance in the cued-recall test accuracy, a generalized linear mixed model was run on the full data. In the model, study decision and study-decision RT were used as independent variables, test accuracy as a dependent variable, and participants and the word-pairs as random effects. Test accuracy was coded as 1 = correct, and 0 = incorrect. Study-decision RT was measured in seconds. The model was significant, $F(2, 2727) = 4.77$, $p = .009$. Study-decision RT was found to account for a significant amount of variance in test accuracy, $F(1, 2727)$, $p = .002$, CI [0.05, 0.21], as shown in Table 2.

Table 2: Results of generalized linear mixed model for the impact of study decision and study-decision response time (RT) on final test accuracies in Experiment 1.

Model	Logit
Fixed part	Coeff. (s.e.)
Intercept	-2.01 (0.17)
Study decision	-0.02 (0.15)
Study-decision RT	0.13 (0.04)
Random part	
Participant	1.20 (1.09)
Swahili-English word-pair	1.08 (1.04)
Deviance (AIC)	13,761.58

Discussion

The correlation results showed that study-decision RT was positively associated with test accuracy. That means, when participants spent longer time (RT) on making study decisions, they were more likely to choose to restudy, they were more likely to have a higher proportion of correct answers in the recall test, and they were more likely to rate higher on their overall learning performance.

The results in the linear mixed model showed study-decision RT could account for a significant amount of variance in the final test accuracy, which re-emphasized the important impact of study-decision RT on test accuracy. Why would learners remember better after spending longer time on making study decisions? They might be guessing or mind-wandering but guessing or mind-wandering would not enhance memory performance. According to the literature, retrieval consumes time (Staszewski, 1988), and contributes to memory and learning performance (Karpicke and Roediger, 2007; Karpicke, 2009; Rawson & Dunlosky, 2011). The positive impact of study-decision RT on test accuracy postulates that the mechanisms underlying study decision making may be an attempted retrieval of the target.

On the other hand, there was no significant association between study decision and test accuracy, indicating the study decisions that were made didn't influence the cued-recall test accuracy even if choosing to restudy could increase the study frequency and duration. The floor effect of test accuracy might have largely limited the potential effects of the cues. However, similar findings are not rare in the literature on both self-paced or experimenter-paced study time (see examples in Nelson & Leonesio, 1988; Zimmerman, 1975). This phenomenon is regarded as labor-in-vain (Nelson & Leonesio, 1988), that extra study time sometimes yields little or no extra gain in subsequent recall. The effect is explained by the notion of attenuation-of-attention (Nelson & Lenonesio, 1988) that the learner pays no more attention to the second repeated presentation. Therefore, our findings are consistent with the literature.

In conclusion, the results showed that the time spent on making study decisions was positively associated with their memory performance. But study decisions, which influenced study duration and frequency, had little impact on later retention. Thus, retrieval is possibly the underlying mechanism that occurs during study decision making and leads to better memory performance. All these suggest learners may monitor their learning processes using retrieval when explicitly cued to control their learning processes (make study decisions).

Experiment 2

The findings in Experiment 1 suggested the possibility that retrieval was involved when learners were asked to make study decisions. Experiment 2 was conducted to further investigate the topic using a within-subject design to have a better control of the study time and possible underlying cognitive activities. The within-subject design was also to increase the statistical power.[1]

Method

Participants Seventy-four college students were recruited and participated the experiment for course credits. None of them participated in Experiment 1.

Stimuli The same stimuli as in Experiment 1 were used in Experiment 2. The number of the Swahili-English word-pairs was reduced to 20 in each of the three conditions to increase overall test accuracy. Thus, each participant learned 60 Swahili English word-pairs (20 items * 3 conditions) in Experiment 2.

Procedure All participants underwent three different learning conditions – single-study condition, retrieval condition, and study-decision condition. A break of 20 seconds was applied between two conditions. The order of the conditions was randomized for each participant. The difficulty levels of all condition were balanced according to the difficulty levels of Swahili-English paired associates published by Nelson and Dunlosky (1994). There were 20 Swahili-English word-pairs in each condition, and 60 items in the entire experiment. The items in each condition were presented in random order. The general procedure of Experiment 2 was similar to that of Experiment 1.

The learning task was the same in each condition: to learn 20 Swahili-English word pairs, but the interventions were different. In the single-study condition, the Swahili-English word-pairs were presented one after another with a 500-millisecond delay in between (no cue). Each item was presented once for 8 seconds. In the retrieval condition, participants saw each item once for 8 seconds, then they were asked to complete a distractor task. After the distractor task, participants were asked to recall the item they just saw and report whether they still remembered it (retrieval cue). The distractor task was the same task used in Experiment 1. In the study-decision intervention, participants were instructed to make a study decision between restudy and advancing to the next item (control cue) after seeing an item for 8 seconds and completing a distractor task. Participants were informed that no additional study time was provided regardless of the study decision. The interventions were self-paced. The amount of written instruction in the retrieval intervention and study-decision intervention were balanced. After seeing all 20 Swahili-English word-pairs in each condition, participants were asked to rate their overall learning performance in that condition on a scale ranging from 0% to 100%. Participants were instructed to complete a cued-recall test in the end of each condition. The cued-recall test was similar to the one used in Experiment 1.

Assumptions of the Underlying Mechanisms

Assumptions of the underlying mechanisms of the retrieval and study-decision interventions were made according to the findings in Experiment 1. When a participant reported that he/she still remembered the Swahili-English word-pair in the retrieval intervention, it was inferred that he/she attempted to retrieve and successfully retrieved the target item (+ +).[2] When a participant reported that he/she didn't remember the pair, it was inferred that he/she attempted to recall but failed to retrieve the word-pair (+ -).[3] The findings from Experiment 1 suggested that learners are likely to practice retrieval when prompted to make a study decision. Therefore, we made the following assumptions: when a participant chose to restudy, it was because that he/she attempted to retrieve the target information but failed (+ -). When a participant chose to advance to the next item immediately, two possible causes were assumed: he/she attempted to retrieve and succeeded (+ +), or he/she didn't even attempt to retrieve the item (- N/A; see Table 3).

Table 3: Assumptions of the cognitive processes underlying the responses in the retrieval and self-study intervention in Experiment 2. Note: "+" refers to the practice of the target cognitive process; "-" refers to a miss practice of the target cognitive process.

Intervention	Responses	Cognitive processes: Retrieval attempt	Cognitive processes: Retrieval success
Retrieval	Remember	+	+
	Not remember	+	-
Study-decision	Next	+	+
		-	N/A
	Study Again	+	-

[1]A small-sample (74 participants), between-subject version of Experiment 2 was conducted and served as a pilot to test whether these results in Experiment 1 could be replicated. In the pilot experiment, participants were randomly divided into four groups: a single-study group, a restudy group, a retrieval group, and a study-decision group, and went through similar procedures as in the corresponding conditions in Experiment 2. No effect of cues was found on test accuracy. The effect size f was conducted from means, $f = .14$. The effect size was small. The statistical power ($1-\beta$) was as low as .15.

Results

A repeated measures ANOVA compared the effects of learning conditions on test accuracies (see descriptive statistics of test accuracies in Table 4). There was no significant condition effect on final test accuracies, $F(2, 146) = .82$, $MSe = .02$, $p = .44$, $\eta^2 = .01$.

Table 4: Means and standard deviations of test accuracies in Experiment 2.

	Test Accuracy		
Condition	*n*	*M*	*SD*
Single-study	74	.37	.21
Retrieval	74	.39	.23
Study-decision	74	.36	.23

In Experiment 2, it took participants 2496.22 ms (*SD* = 770.10) on average to report their retrieval results when learning in the retrieval condition, and it took participants averagely 2248.51 ms (*SD* = 868.83) to make a study decision when learning in the study-decision condition. A paired sample t-test showed significant difference in the intervention response time between the retrieval and study-decision conditions, $t(73) = -2.71$, $p < .01$, indicating that participants spent more time on reporting the retrieval results than on making study decisions.

A paired sample t-test found within-subject effect of responses on test accuracies: participants achieved higher test accuracies with the items that they successfully retrieved ($M = .42$, $SD = .24$), than the items that they failed to retrieve ($M = .17$, $SD = .29$) in retrieval intervention, $t(59) = 6.92$, $p < .001$. Participants achieved higher test accuracies with the items that they decided to have no restudy ($M = .40$, $SD = .26$), $t(49)=4.79$, $p < 0.001$ than the items that they decided to restudy ($M = .21$, $SD = .25$), as in Figure 1.

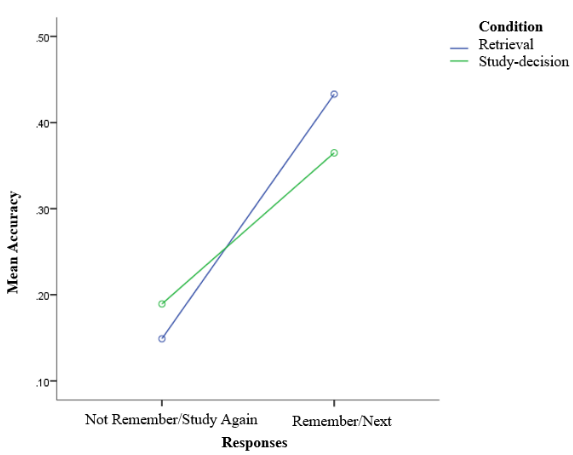

Figure 1: Within-subject effect of intervention response on test accuracies in the retrieval and study-decision conditions in Experiment 2.

Discussion

The results in Experiment 2 successfully replicated the findings from Experiment 1 that the retrieval-based metacognitive monitoring is likely to be performed during study-decision making.

The results showed the retrieval group was more likely to successfully recall the items they reported as remembered over the items they reported as forgotten in the cued-recall test. The study-decision group was more likely to successfully recall the items they chose no-restudy over the items they chose to restudy. The findings indicate participants were unlikely to make study decisions at random. What's more, the cognitive activities underlying study-decision making in the study-decision group seemed to resemble the cognitive activities underlying retrieval-result in the retrieval group. The observed significant within-subject effects of responses on accuracy in retrieval and study-decision conditions also helped confirm the assumptions of mechanisms underlying interventions. Restudy decisions ("study again") may be resulted from an unsuccessful attempt of retrieval, while no-restudy decisions ("next") may be resulted from a successful retrieval or no attempt of retrieval (as in Table 3). Thus, the hypothesis was supported that retrieval is likely to be practiced when learners are cued to control (i.e., make study decisions).

However, the statistical analysis also revealed that reporting retrieval results took more time than making study decisions. Since the prior findings indicate that learners attempted retrieval in both retrieval and study-decision interventions, the difference between the retrieval RT and study-decision RT might be resulted from a different retrieval fashion underlying the two interventions. Our explanation is, learners might attempt retrieval more often or at a deeper level when explicitly asked to retrieve and report the retrieval results, which lead to longer overall response time. Thus, when learners are cued to control (e.g., make study decisions), retrieval is likely to be practiced at a lower frequency or at a more shallow level (e.g., feeling of knowing instead of free recall results) than when learners are cued to retrieve.

Unexpectedly, Experiment 2 also confirmed that the cues (no cue, retrieval cues, and study-decision cues) don't affect later retention. The results suggest retrieval practice may not affect learning outcomes in the current learning setting, regardless whether it is cued explicitly.

Conclusion

The current study suggest that learners are likely to monitor their learning processes by using retrieval when they are explicitly cued to control, but in an inconsistent fashion in self-regulated learning. Theoretically, the current study indicates that retrieval may function as a monitoring tool when learners are cued to control their learning processes. From a practical perspective, the current study suggests presenting learning items without pause may be an efficient presentation strategy. The analysis was primarily based on behavioral data (e.g., decision making response time, and final test accuracy), which suggests the possibility of measuring retrieval-based metacognitive monitoring using objective and online, but indirect methods.

The current study has limitations. Firstly, the effect of a study-decision cue is not clear. The study-decision cue, which was used as a control cue in this study was expected to prompt study-decision making only, but it is argued that the study decision cue might also be indirectly treated as a monitoring cue. Secondly, metacognitive monitoring in this study was measured and inferred through the study-decision response time, and final recall accuracy. These measures are objective and online, but they are not direct measures of monitoring, and rely heavily on interpretation. Eye-tracking, as proposed by Veenman (2005), has a potential of providing objective, online, and direct evidence of cognitive processes, including metacognitive processes. It may be used to capture monitoring processes more directly, which can be followed up by future studies.

Acknowledgements

Special thanks to the EPSY lab, Dr. Christopher A. Was, Dr. Bradley J. Morris, and Ms. Rachael Todaro for their support and assistance.

References

Bryce, D., Whitebread, D., & Szűcs, D. (2014). The relationships among executive functions, metacognitive skills and educational achievement in 5 and 7-year-old children. *Metacognition and Learning, 10*(2), 181-198.

Dunlosky, J., Hertzog, C., Kennedy, M. R. F., & Thiede, K. W. (2005). The self-monitoring approach for effective learning. *International Journal of Cognitive Technology, 10*, 4–11.

Dunlosky, J., & Metcalfe, J. (2009). *Metacognition.* Thousand Oaks, CA: Sage.

Dunlosky, J., & Nelson, T. O. (1992). Importance of the kind of cue for judgments of learning (JOL) and the delayed-JOL effect. *Memory & Cognition, 20*, 374–380.

Karpicke, J. D. (2009). Metacognitive control and strategy selection: Deciding to practice retrieval during learning. *Journal of Experimental Psychology: General, 138*, 469–486.

Karpicke, J. D., & Roediger, L. (2007). Repeated retrieval during learning is the key to long-term retention. *Journal of Memory and Language, 57*, 151–162.

Koriat, A. (2007). Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *Cambridge handbook of consciousness*. New York: Cambridge University Press.

Landauer, T. K., & Bjork, R. A. (1978). Optimum rehearsal patterns and name learning. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory*. London: Academic Press.

Metcalfe, J., & Finn, B. (2008). Familiarity and retrieval processes in delayed judgments of learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(5), 1084.

Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The “delayed-JOL effect”. *Psychological Science, 2*(4), 267-270.

Nelson, T. O., & Dunlosky, J. (1994). Norms of paired associate recall during multitrial learning of Swahili-English translation equivalents. Memory, 2(3), 325-335.

Nelson, T. O., & Leonesio, R. J. (1988). Allocation of self-paced study time and the" labor-in-vain effect.". *Journal of experimental psychology: Learning, Memory, and Cognition, 14*(4), 676.

Nelson, T.O., & Narens, L. (1990). Metamemory: A theoretical framework and some new findings. In G.H. Bower (Ed.), *The psychology of learning and motivation*. San Diego: Academic Press.

Nietfeld, J., Cao, L., & Osborne, J. (2005). Metacognitive Monitoring Accuracy and Student Performance in the Postsecondary Classroom. The Journal of Experimental Education, 74(1), 7-28.

Rawson, K. A., & Dunlosky, J. (2011). Optimizing schedules of retrieval practice for durable and efficient learning: How much is enough? *Journal of Experimental Psychology: General, 140*, 283–302.

Renner, C. H., & Renner, M. J. (2001). But I thought I knew that: using confidence estimation as a debiasing technique to improve classroom performance. *Applied Cognitive Psychology, 15*, 23–32.

Roebers, C. M., Krebs, S. S., & Roderer, T. (2014). Metacognitive monitoring and control in elementary school children: Their interrelations and their role for test performance. Learning and Individual Differences, 29, 141–149.

Staszewski, J. J. (1988). Skilled memory and expert mental calculation. In M. T. H. Chi, R. Glaser, & M. J. Farr (Eds.), *The nature of expertise.* Hillsdale, NJ: Erlbaum.

Thiede, K. W., Anderson, M. C. M., & Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of Educational Psychology, 95*(1), 66-73.

Veenman, M. V. J. (2005). The assessment of metacognitive skills: What can be learned from multimethod designs? In C. Artelt & B. Moschner (Eds.), Lernstrategien und Metacognition: Implikationen für Forschung und Praxis (pp. 75–97). Berlin: Waxmann.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. Graesser (Eds.), *Metacognition in educational theory and practice*. Mahwah: Erlbaum.

Zimmerman, J. (1975). Free recall after self-paced study: A test of the attention explanation of the spacing effect. *American Journal of Psychology, 88*, 227-291.

Zimmerman, B. J., & Kitsantas, A. (1999). Acquiring writing revision skill: Shifting from process to outcome self-regulatory goals. *Journal of educational Psychology, 91*, 1-10.