


Testing the Effectiveness of Retrieval-Based Learning in Naturalistic School Settings

SAGE Open
October-December 2021: 1–7
© The Author(s) 2021
DOI: 10.1177/21582440211061569
journals.sagepub.com/home/sgo


Juana M. Ortega-Tudela¹ , María-Teresa Lechuga¹,
Miriam Bermúdez-Sierra¹, and Carlos J. Gómez-Ariza¹

Abstract

While the learning benefits of retrieval activities have been clearly demonstrated in laboratory settings, evidence on their usefulness in naturalistic school settings is still scant. The goal of the present studies was to investigate the feasibility and effectiveness of retrieval-based learning in children (fourth and sixth grades) when school teachers themselves design and implement retrieval activities relating to genuine curriculum contents. Three studies were conducted in a public elementary school with fourth and sixth graders and their teachers. Two of the studies involved mathematics and one dealt with social sciences. Teachers used learning activities that required students to recall part of previously taught concepts, while different concepts in the same unit were worked through with those learning activities that were normally used by each teacher. Two out of three studies revealed that, relative to business-as-usual learning activities, performing retrieval activities during classes led to better performance in the assessments at the end of the lessons. Overall, our finding provides preliminary evidence that retrieval activities can enhance learning in elementary school children when they are devised by teachers in the exercise of their professional duties. These results have important practical implications and suggest that, if teachers are aware of the value of retrieval activities in fostering meaningful learning, these activities could be successfully embedded in their daily duties even when considering the constraints imposed by school reality.

Keywords

knowledge acquisition, learning activities, learning strategies, retrieval-based learning, teachers' innovation

Introduction

At the current time, we are witnessing a renewed enthusiasm for integrating Cognitive Science and Education, with an emphasis on evidence-based practices (Cook et al., 2012). However, even with the consideration that educational practices are continually changing, some of these changes are not supported by scientific evidence. As Roediger (2013) has recently emphasized, translational educational science should be adopted by schools and educators, to improve education, on the basis of the latest cognitive and educational research findings.

In this context, one cutting-edge topic bridging cognitive science and educational practice is that of retrieval-based learning (Nunes & Karpicke, 2015). In essence, retrieval-based learning requires students to actively engage in activities that rely on recalling previously studied information (retrieval practice; i.e., responding to a teacher's questions by using a clicker), with this practice usually leading to enhanced performance on subsequent learning tests (the so-called "testing effect") by comparison with other traditional

learning activities such as rereading or concept mapping (i.e., Karpicke & Blunt, 2011; Lechuga et al., 2015; Ortega-Tudela et al., 2019). Indeed, it is now well established that retrieval (and testing) can play a role as a promoter of meaningful learning, and retrieval activities have now been proven to benefit learning of academic content and to facilitate transfer to related concepts (see Karpicke, 2017, for a review). To date, however, most research on retrieval-based learning has derived from laboratory settings, with few efforts made as yet to examine whether the benefits of retrieval practice are also observable when students deal with their own course materials in real classroom environments (Fazio & Marsh, 2019). **The main aim of the present work was to contribute in this regard by investigating the**

¹University of Jaén, Spain

Corresponding Author:

Juana M. Ortega-Tudela, Department of Pedagogy, University of Jaén, Campus Las Lagunillas s/n, 23071 Jaén, Spain.
Email: jmortega@ujaen.es



effectiveness of retrieval-practice activities in authentic classroom environments.

From Laboratory to Naturalistic Learning Environments in Primary Education

The lack of studies on retrieval-based learning in naturalistic settings is remarkable, especially in the case of primary education, most likely due to the complexity of its educational methodologies and the difficulty of getting teachers actively involved in the implementation of “new” learning activities with which they are not familiar (Cook et al., 2012). Thus, while the benefits of retrieval-based learning have been shown for both adolescents and adults in laboratory classroom settings (Karpicke, 2017), much less research has been directed at studying the potential benefits for young children (Fazio & Marsh, 2019; Karpicke et al., 2016; Moreira et al., 2019).

More relevant from an applied standpoint, there remain only a few studies that have focused on primary and middle-school students and their learning of actual curriculum topics (rather than artificial materials), such as history facts (Carpenter et al., 2009), science and geography topics (Karpicke et al., 2014), general knowledge topics (Marsh et al., 2012), and vocabulary terms (Goossens et al., 2014). To give an example of this, McDaniel et al. (2013) showed in middle-school students that learning of science concepts improved (as measured by class summative assessments) for content that had previously been tested (providing feedback) using multiple-choice quizzes (a retrieval-practice activity), relative to concepts that had not been tested. In brief, while there are strong reasons to expect that elementary children may show retrieval practice effects, evidence about its usefulness in promoting learning in real classes is still limited. To the best of our knowledge, in all these studies the retrieval activities were created and implemented by experimenters who were guests in the classrooms, such that the role of the practicing teachers as active agents for learning was conveniently minimized.

Thus, before decisively claiming that retrieval practice should be somehow incorporated in elementary education classrooms as a learning activity, it seems reasonable to ensure that retrieval-based learning is possible under the constraints imposed by the school reality. To this end, we embraced the perspective of effectiveness studies as described by Green et al. (2019), which examine whether significant real-world impact is observed when an intervention is used in less than ideally controlled settings. This approach holds special relevance in light of the arguments put forward by Roediger (2013), who highlighted the need to improve education practices on the basis of the latest cognitive and educational research findings. Hence, the main aim of the present research was to investigate the effectiveness of retrieval-practice activities in children when they are

embedded in authentic classroom environments and are directed by practicing teachers dealing with genuine academic contents.

The Current Studies

We report three independent studies in which selected classes teachers used learning activities that required students to recall part of previously taught concepts, while different concepts in the same unit were worked through with other learning activities (those that were normally used by the teacher). Based on previous evidence, our general hypothesis was that children would exhibit better learning for those concepts that had been the target of retrieval practice relative to those that have been worked using other techniques that are not based largely on retrieval.

Importantly, and in contrast to previous studies, elementary school teachers planned and developed the learning activities by embedding them into the daily dynamics of their classes. While this naturally led to differences across the studies (they vary according to the number of retrieval activities performed by the teacher, whether or not feedback was provided after the retrieval activity, and the characteristics of the specific activity being performed), these differences would provide us with the opportunity to assess the effectiveness of teacher-directed retrieval-based learning (Green et al., 2019). The researchers supervised the process and cleared up in advance any doubts that teachers had regarding the design and implementation of the retrieval activities, while interfering as little as possible in the development of the classes. To our knowledge, no previous study on retrieval-based learning in school settings has taken this approach.

General Methods

Because the three studies share a good number of features, we first report their commonalities and then specify the details for each.

Participants

Altogether, 70 children who were enrolled in different grades (fourth and sixth grade; ages from 9 to 11 years old) participated in the studies along with their actual (different across studies) teachers, who voluntarily joined the project after receiving a brief training course on retrieval-based learning. The number of children in each class ranged from 13 to 19 ($M = 17.6$, $SD = 5.03$). Two of the studies involved only one class, whereas the third one was carried out with two different classes of sixth-grade students. This fact was a consequence of the duties of the participating teachers.

The studies were conducted in a public elementary school located in a middle-class neighborhood. All participants were Castilian Spanish speakers and monolinguals. None of them had specific learning difficulties. Importantly, parents

were informed through the school board and gave written informed consent in advance. Also, the university's research ethics committee approved the studies.

The teachers were experienced professionals (more than 5 years of teaching experience) who had been in charge of at least one group of students since the start of the school year. They received three 45-minute sessions of instruction in which one of the experimenters showed them (a) basic findings (with children and adults) from the experimental literature on the testing effect, (b) the most relevant theoretical interpretations of the benefits of retrieval-based activities, and (c) how to create and make use of these activities through adapting them to their classroom materials and dynamics.

Materials and Procedure

Two of the studies involved mathematics and one dealt with social sciences. In all of them, teachers made use of learning activities that required children to recall part of previously taught concepts, while other concepts from the same unit were worked through with those learning activities that were normally used by the teacher. Hence, the three studies comprised a within-subject design regarding learning activities.

All of the groups worked in their lessons using their textbooks, which included different materials and activities to facilitate learning (i.e., concept explanations, exercises, and examples). The teachers taught their lessons as they usually would, without introducing any changes into the dynamic of their classes, except for the inclusion of brief retrieval tasks at a certain point. The content (i.e., concepts, facts) of the unit/lesson to be either retrieved or worked through using ordinary activities was chosen by the teachers, who also made the decisions on which specific activities were to be used in addition to the right moment to do so.

The retrieval activities consisted of questions for which students were required to write down their answers on paper, without having access to the texts. Some (open) questions were to be answered with one word or several words (Studies 1 and 2), while others were multiple option questions (Study 3). In all cases, there was only one correct answer to the question. The teachers scored the participants' responses.

Learning was assessed once the thematic unit was finished, although the precise time that elapsed between the end of the lesson and the learning test varied across studies depending upon the teachers' requirements. The learning tests were created by the teachers in collaboration with the researchers. An effort was made to design the learning tests so that their questions differed in nature from those used in the retrieval activities. This was done to prevent an eventual advantage for the retrieved contents (relative to non-retrieved ones) due to the simple repetition of questions (and answers). Importantly, the learning tests included questions on retrieved and non-retrieved concepts from the lessons. The students were allowed to take the time they needed to answer the set

of questions, which came to 15 minutes on average. The teachers elaborated templates to score student responses so that the scoring procedure was relatively simple and free of ambiguity. Responses were scored by a PhD student who was blind to the conditions governing the learning tests.

Study I. Mathematics, Grade 4

Method

Participants. One classroom group with 13 fourth graders (ten girls, $M_{age}=9.9$, $SD_{age}=0.3$) took part in this study, although only ten students were included in the analysis because three of the students failed to attend at least one session.

Materials and procedure. The material was selected from the geometric objects unit included in the fourth-grade math book for elementary school. The teacher chose ten relevant concepts/facts about geometric objects before starting the lesson. The difficulty of the concepts was similar, based on the teacher's experience. Five of them were randomly selected to be the target of retrieval activities. The teacher also decided that retrieval activities would consist of answering short questions on concepts worked on in previous classes. In each retrieval activity, one question was provided on a piece of paper on which students wrote/drew their responses, which could be one or two words, a number, or a simple shape. The teacher provided feedback on correct answers immediately. The remaining five target concepts of the unit were not worked on using retrieval activities. Instead, they were reviewed with activities such as rereading, summarization or problem solving that the teacher usually used in their classes.

The class had five 45-minute lectures of mathematics per week and, as scheduled, the teacher spent 2 weeks dealing with the geometric shapes unit. Over the course of the 2 weeks, one retrieval activity or non-retrieval activity (focused on concepts that had been taught on previous days) was carried out at the beginning of a class, alternating between the two types of activity from day to day.

One week after finishing the thematic unit, the students took a learning test that included seven multiple-choice (three alternative responses) questions: four of them referred to concepts reviewed/worked through using retrieval activities and the remaining three referred to concepts reviewed/worked through using conventional activities. The teachers developed templates to score the students' responses. Only one of the three possible response choices was considered correct.

Results

In the present and following studies, the effect of retrieval activities on learning was examined using an analysis of variance (ANOVA) on the proportion of correct responses in

Table 1. Accuracy (Means and Standard Deviations) in the Retrieval Activities (During Learning) and in the Learning Tests (At the End of the Lessons) as a Function of Learning Activities and Studies.

	Retrieval activity	Learning activity	
		Retrieval-based	Business-as-usual
Study 1 (n = 10)	.81 (.21)	.93 (.12)	.53 (.17)
Study 2 (n = 18)	.77 (.22)	.90 (.15)	.76 (.13)
Study 3 (n = 32)	.74 (.20)	.39 (.27)	.35 (.24)

the learning test, with the type of learning activity (retrieval vs. non-retrieval) as the factor. In addition, we accomplished non-parametric (Spearman's rho) correlation analyses to explore if accuracy during the retrieval activities were predictive of performance in the learning test.

The repeated measures ANOVA¹ showed that there was a reliable effect of learning activity, $F(1, 9) = 50.46$, $Mse = 0.02$, $p < .001$, $p\eta^2 = .85$ (see Table 1). The students produced more correct responses for concepts that had been worked through with retrieval activities ($M = 0.93$; $SD = 0.12$) than for concepts that had been worked through with alternative activities ($M = 0.53$; $SD = 0.17$).

Correlation analyses failed to show a reliable association between the proportion of correct responses in the retrieval activities ($M = 0.81$; $SD = 0.21$) and performance in the learning test ($ps > 0.40$).

Study 2. Mathematics, Grade 6

Method

Participants. The participants comprised one classroom group of nineteen sixth graders (nine girls, $M_{age} = 12.16$, $SD_{age} = 0.51$). One girl did not attend the final learning test.

Materials and procedure. The material from this study was selected from the sixth-grade math book for elementary school. The topics evaluated in this study were geometric objects, Platonic solids and capacity units of measure. Over the course of 2 weeks, the students had a 45-minute lecture on each of the 5 days of the school week and used the above-mentioned textbook. The teacher chose ten relevant concepts at the beginning of the thematic unit and they explained the lesson as they usually would do. Five of the concepts were randomly selected to be the target of retrieval activities.

Retrieval activities consisted of answering three paper and pencil questions regarding concepts taught in previous classes. Only two retrieval-activity sessions were carried out (at the beginning of the third and seventh classes). In each retrieval activity, three questions were provided on a piece of paper on which students wrote/drew their responses, which could be one or two words, a number, or a shape. The teacher provided feedback on correct answers immediately. The remaining five target concepts of the unit were worked

through with rereading, summarization and problem-solving activities that the teacher usually used in their classes.

The final learning test took place after finishing the unit and consisted of 10 multiple-choice questions on the target concepts (four related to retrieved concepts and six referring to non-retrieved concepts). The learning test involved the use of Plickers,² since the teacher usually employed this tech tool.

Results

The repeated measures ANOVA on the proportion of correct responses in the learning test showed a reliable effect of learning activity, $F(1, 17) = 7.55$, $Mse = 0.025$, $p = .01$, $p\eta^2 = .74$. Questions on concepts that were worked on in the classroom using retrieval activities received more correct responses ($M = 0.90$; $SD = 0.15$) than questions on concepts that were addressed using other learning activities ($M = 0.76$; $SD = 0.13$) (see Table 1).

In addition, accuracy in the retrieval activities ($M = 0.77$; $SD = 0.22$) predicted performance in the learning test. Specifically, the rate of correct responses during retrieval practice correlated with accuracy in responding to questions on concepts that had been worked through with retrieval activities ($\rho = .54$, $p = .02$). No association emerged regarding concepts that had been worked through with non-retrieval activities ($p > .80$).

Study 3. Social Sciences, Grade 6

Method

Participants. Participants were 38 students from two classroom groups (19 children in each, with even gender distribution; 17 girls) in the sixth grade ($M_{age} = 12.08$, $SD_{age} = 0.42$). Four participants were eliminated from the analyses because they failed to attend at least one of the learning sessions.

Materials and procedure. The learning material (the unit "Spain in the XIX century") was selected from the social science textbook. The students had two 45-minute lectures of social science per week, and the teacher taught the selected unit in the usual way over 6 weeks, using the above-mentioned textbook. In advance, the teacher selected sixteen topics from the unit so that half of them could be worked on using a retrieval activity.

The teacher decided to implement a retrieval activity that consisted of a paper and pencil test composed of eight multiple-choice questions, without providing feedback on the responses. Importantly, since the teacher was in charge of two class groups, the concepts to be worked through using retrieval and non-retrieval activities were counterbalanced across groups. Hence, at the end of the thematic unit the group A engaged in the retrieval activity for half of the concepts whereas the group B did retrieval of the remaining eight target concepts of the unit.

One week after finishing the unit, students took a final learning test on the target concepts, consisting of 16 short-answer questions. The test was identical for both groups and students were told to take the time they needed to answer the questions, which came to approximately 30 minutes on average.

Results

The repeated measures ANOVA on the proportion of correct responses in the final learning test did not reveal differences between the students' performance for retrieved ($M=0.39$; $SD=0.27$) and non-retrieved concepts ($M=0.35$; $SD=0.24$), $F(1, 31)=1.73$, $Mse=0.016$, $p=.198$, $p\eta^2=.05$ (see Table 1). However, accuracy during the retrieval activities ($M=0.74$; $SD=0.20$) predicted success at responding questions on both type of concepts (retrieved concepts: $\rho=.52$, $p<.001$; non-retrieved concepts: $\rho=.62$, $p<.001$).

General Discussion

While it is now well established that activities that rely on retrieving information from memory promote meaningful learning of this information and facilitate transfer to related concepts (see Karpicke, 2017, for a review), thus far research on retrieval-based learning has mostly come from laboratory studies with adult participants. Importantly, the few studies that have focused on curricular contents have also demonstrated the learning benefits of retrieval activities in elementary and middle-school students (i.e., Carpenter et al., 2009; Karpicke et al., 2014; Marsh et al., 2012), although in all these studies the learning activities were designed and implemented by the experimenters, with the students' actual teachers playing no relevant role in the process. The present research aimed to go a step further in investigating the effectiveness of retrieval-based learning in primary schools by engaging practicing teachers in the conception, design, and implementation of the retrieval activities, which they incorporated into their otherwise typical classes.

Studies 1 and 2 revealed that, relative to business-as-usual learning tasks, having students perform retrieval activities during the classes led to better performance (a testing effect) in assessments at the of end lessons. This finding is remarkable because (a) it is in line with the results of a number of well-controlled experiments showing that practicing retrieval

may be a powerful way to enhance learning (for a review, see Karpicke, 2017), and (b) it demonstrates for the first time that retrieval-based learning is feasible (and may be effective) even when it is managed by real teachers in authentic classroom environments. Thus, while the present studies conveniently tempered the experimental control that is necessary in other types of research, their results are of special relevance to discovering the feasibility and effectiveness of implementing retrieval activities as a way of enhancing learning in elementary school children even when such activities are devised by teachers in the exercise of their professional duties.

Contrary to Studies 1 and 2, Study 3 failed to show learning differences. Having children engage in a task that required them to recall some relevant concepts of the lesson did not lead to better performance in the final assessment than typical procedures. While the reasons behind this null result are not obvious to us, [it should be noted that simply judging by the overall low performance (about 37% on average) in the assessment, one could argue that the relative stringency of the final test might have made it difficult to reveal significant differences], a few features of Study 3 deserve attention here. First of all, a potentially relevant difference from Studies 1 and 2 is that in Study 3 the retrieval-based learning activity consisted of a multiple-choice test without feedback. Whether or not a multiple-choice test may be effective as a learning tool is still unclear (i.e., Greving & Richter, 2018; Moreira et al., 2019). On the one hand, the act of selecting the correct response from several presented options is thought to demand less retrieval effort than generating a response from memory and has been shown to produce limited benefits (i.e., Kang et al., 2007). On the other hand, multiple-choice tests necessarily expose students to plausible incorrect responses that could cause confusion and unsuccessful learning (Marsh et al., 2012). Additionally, feedback was not provided to children during retrieval practice in Study 3, which may have played a role in buffering the benefit of retrieval activities. Though the role of corrective feedback in retrieval-based learning is still poorly understood, in a recent review Moreira et al. (2019) concluded that the use of retrieval practice in classroom settings may be effective regardless of whether feedback is provided or not, there is evidence that feedback may be especially useful in elementary school children (i.e., Lipko-Speed et al., 2014). In any case, it is important to underline here that, such as is usually the case in real school environments, it was the teacher in charge of each classroom who made the decisions on how to proceed with the learning (including retrieval activities that they used in their classes (format, feedback, content, etc.), which demonstrates that retrieval-based learning may be feasible and effective in naturalistic school settings.

It must be highlighted (and recognized) here that participating teachers played a primary role in the present studies. They spent time attending the instruction sessions on

retrieval-based learning, but they also led the process of adaptation and implementation of the (retrieval and non-retrieval) activities to test their relative effectiveness. While the experimenters advised them and tried to answer their procedural questions, decisions were made by the teachers themselves, who incorporated the new activities into their classroom dynamics. To our knowledge, no study to date has explored the benefits of retrieval-based activities in naturalistic school settings, bringing the leading role of real teachers into focus.

Practical Implications

Given the overall positive results derived from Studies 1 and 2, it seems reasonable to posit that learning activities that strongly rely on retrieval can be created and used (in addition to and/or in combination with other activities such as concept mapping; i.e., Ortega-Tudela et al., 2019) by teachers in their classes in order to enhance conceptual learning in children. After explaining concepts and facts for the first time, teachers typically ensure students' learning by using a variety of activities such as rereading, concept mapping, summarizing, or homework setting. However, they are usually reluctant to engage students in activities that strongly rely on retrieval. Neither students (children or adults) nor teachers usually consider retrieval practice to substantially contribute to learning (Tullis et al., 2013). Instead, retrieval is uniquely seen as a component of assessment, and tests are thought to serve only to measure knowledge. Hence, the incorporation of testing procedures as learning tools in the classrooms would necessarily require better communication flows between cognitive/education researchers and education practitioners in order to allow teachers to become familiar with such activities and their potential impact on learning. In addition to the direct benefits that retrieval activities have been shown to have (for a review, see Karpicke, 2017), testing may also contribute to learning in more indirect ways (i.e., it provides students with more and better opportunities to detect possible gaps in their knowledge; Roediger et al., 2011).

Limitations and Future Directions

Although our findings may be interpreted with optimism, we recognize the limitations of the present studies. The small number of students, teachers and courses participating, and the short-term assessment of learning that was considered, force us to be cautious regarding our results. Hence, further research is warranted to replicate the present positive findings as well as to shed light on the conditions that can modulate retrieval-based learning in natural environments. Thus, for example, it will be of relevance to determine the role of corrective feedback in elementary school children, which has been proposed to be especially useful when it comes to learning with retrieval activities (i.e., Lipko-Speed et al., 2014).

In our view, the naturalistic approach that was taken in the present studies is necessary, as a forward step, to endorse the use of retrieval-based activities in natural learning environments. The considerable diversity and complexity of educational settings make it difficult to directly generalize findings from basic research to real classroom dynamics. Thus, the direct involvement of teachers in this type of research will provide valuable findings on human learning and help design better interventions in genuine educational settings. Practitioners of education need to know about the potential benefits of testing as a tool of learning (not only as a tool of assessment), and we hope that research initiatives like the one presented here will contribute to raise students' and teachers' awareness of retrieval-based strategies.

Conclusion

Our results provide preliminary evidence that practicing retrieval may be an effective learning strategy in naturalistic school settings. Also, these results highlight the important role of teachers in implementing retrieval-based strategies in their daily duties. We expect these findings to foster further research of this nature in order to define with greater precision the extent to which retrieval-based learning is feasible and effective in children's elementary education.

Acknowledgments

We thank Ana Martos, Libertad López, Antonia Limones and Manuel Rosa for their support and implication in this project. We also thank the rest of the staff and children from the school CEIP Alcalá Venceslada of Jaén (Spain) for their help.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was funded by Consejería de Economía, Conocimiento, Empresas y Universidades from Junta de Andalucía, Spain. Programa Operativo FEDER Andalucía 2014-2020 (Proyectos_FEDER_1262768).

ORCID iD

Juana M. Ortega-Tudela  <https://orcid.org/0000-0003-3649-3524>

Notes

1. Due to the small number of participants in this study, we also performed a (non-parametric) Wilcoxon signed-rank test that confirmed that the retrieval activities led to better performance ($p = .005$) than the typical learning activities.
2. Plickers is a free tech tool that allows teachers to collect data without having students use devices or paper and pencil. Students have a card with a QR code printed upon it; each

edge of this QR code has a letter (a, b, c, or d) that corresponds to different response options. They simply need to rotate their card to show their answer. Teachers need to create the questions on the application first. Then, in the classroom, the teacher asks the questions to the students, who respond by using their cards. Finally, the teacher scans the cards using a mobile phone or a tablet. Plickers automatically scores the students' responses and provides feedback on performance.

References

- Carpenter, S. K., Pashler, H., & Cepeda, N. J. (2009). Using tests to enhance 8th grade students' retention of US history facts. *Applied Cognitive Psychology, 23*(6), 760–771. <https://doi.org/10.1002/acp.1507>
- Cook, B. G., Smith, G. J., & Tankersley, M. (2012). Evidence-based practices in education. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, & J. Sweller (Eds.), *APA educational psychology handbook, vol. 1. Theories, constructs, and critical issues* (pp. 495–527). American Psychological Association.
- Fazio, L. K., & Marsh, E. J. (2019). Retrieval-based learning in children. *Current Directions in Psychological Science, 28*(2), 111–116. <https://doi.org/10.1177/0963721418806673>
- Goossens, N. A. M. C., Camp, G., Verkoeijen, P. P. J. L., & Tabbers, H. K. (2014). The effect of retrieval practice in primary school vocabulary learning. *Applied Cognitive Psychology, 28*(1), 135–142. <https://doi.org/10.1002/acp.2956>
- Green, S. C., Bavelier, D., Kramer, A. F., Vinogradov, S., Anson, U., Ball, K. K., Bingel, U., Chein, J. M., Colzato, L. S., Edwards, J. D., Facoetti, A., Gazzaley, A., Gathercole, S. E., Ghisletta, P., Gori, S., Granic, I., Hillman, C. H., Hommel, B., Jaeggi, S. M., ... Witt, C. M. (2019). Improving methodological standards in behavioral interventions for cognitive enhancement. *Journal of Cognitive Enhancement, 3*, 2–29. <https://doi.org/10.1007/s41465-018-0115-y>
- Greving, S., & Richter, T. (2018). Examining the testing effect in university teaching: Retrieval and question format matter. *Frontiers in Psychology, 9*, 2412. <https://doi.org/10.3389/fpsyg.2018.02412>
- Kang, S. H. K., McDermott, K. B., & Roediger, H. L., III. (2007). Test format and corrective feedback modify the effect of testing on long-term retention. *The European Journal of Cognitive Psychology, 19*(4–5), 528–558. <https://doi.org/10.1080/09541440601056620>
- Karpicke, J. D. (2017). Retrieval-based learning: A decade of progress. In J. H. Byrne (Ed.), *Learning and memory: A comprehensive reference* (2nd ed., pp. 487–514). Academic Press.
- Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science, 331*(6018), 772–775. <https://doi.org/10.1126/science.1199327>
- Karpicke, J. D., Blunt, J. R., & Smith, M. A. (2016). Retrieval-based learning: Positive effects of retrieval practice in elementary school children. *Frontiers in Psychology, 7*, 350. <https://doi.org/10.3389/fpsyg.2016.00350>
- Karpicke, J. D., Blunt, J. R., Smith, M. A., & Karpicke, S. S. (2014). Retrieval-based learning: The need for guided retrieval in elementary school children. *Journal of Applied Research in Memory and Cognition, 3*(3), 198–206. <https://doi.org/10.1016/j.jarmac.2014.07.008>
- Lechuga, M. T., Ortega-Tudela, J. M., & Gómez-Ariza, C. J. (2015). Further evidence that concept mapping is not better than repeated retrieval as a tool for learning from texts. *Learning and Instruction, 40*, 61–68. <https://doi.org/10.1016/j.learninstruc.2015.08.002>
- Lipko-Speed, A., Dunlosky, J., & Rawson, K. A. (2014). Does testing with feedback help grade-school children learn key concepts in science? *Journal of Applied Research in Memory and Cognition, 3*(3), 171–176. <https://doi.org/10.1016/j.jarmac.2014.04.002>
- Marsh, E. J., Fazio, L. K., & Goswick, A. E. (2012). Memorial consequences of testing school-aged children. *Memory, 20*(8), 899–906. <https://doi.org/10.1080/09658211.2012.708757>
- McDaniel, M. A., Thomas, R. C., Agarwal, P. K., McDermott, K. B., & Roediger, H. L. (2013). Quizzing in middle-school science: Successful transfer performance on classroom exams. *Applied Cognitive Psychology, 27*(3), 360–372. <https://doi.org/10.1002/acp.2914>
- Moreira, B. F. T., Pinto, T. S. S., Starling, D. S. V., & Jaeger, A. (2019). Retrieval practice in classroom settings: A review of applied research. *Frontiers in Education, 4*, 5. <https://doi.org/10.3389/educ.2019.00005>
- Nunes, L. D., & Karpicke, J. D. (2015). Retrieval-based learning: Research at the interface between cognitive science and education. In R. A. Scott & S. M. Kosslyn (Eds.), *Emerging trends in the social and behavioral sciences: An interdisciplinary, searchable, and linkable resource* (pp. 1–16). John Wiley & Sons, Inc.
- Ortega-Tudela, J. M., Lechuga, M. T., & Gómez-Ariza, C. J. (2019). A specific benefit of retrieval-based concept mapping to enhance learning from texts. *Instructional Science, 47*(2), 239–255. <https://doi.org/10.1007/s11251-018-9476-y>
- Roediger, H. L., III. (2013). Applying cognitive psychology to education: Translational educational science. *Psychological Science in the Public Interest, 14*(1), 1–3. <https://doi.org/10.1177/1529100612454415>
- Roediger, H. L., III, Putnam, A. L., & Smith, M. A., III. (2011). Ten benefits of testing and their applications to educational practice. In J. P. Mestre & B. H. Ross (Eds.), *The psychology of learning and motivation: Vol. 55. The psychology of learning and motivation: Cognition in education* (pp. 1–36). Elsevier Academic Press.
- Tullis, J. G., Finley, J. R., & Benjamin, A. S. (2013). Metacognition of the testing effect: Guiding learners to predict the benefits of retrieval. *Memory & Cognition, 41*, 429–442. <https://doi.org/10.3758/s13421-012-0274-5>