Educ Psychol Rev (2012) 24:437–448
DOI 10.1007/s10648-012-9210-2

REVIEW ARTICLE

# The Value of Applied Research: Retrieval Practice Improves Classroom Learning and Recommendations from a Teacher, a Principal, and a Scientist

**Pooja K. Agarwal · Patrice M. Bain · Roger W. Chamberlain**

Published online: 19 August 2012


**Abstract** Over the course of a 5-year applied research project with more than 1,400 middle school students, evidence from a number of studies revealed that retrieval practice in authentic classroom settings improves long-term learning (Agarwal *et al.* 2009; McDaniel *et al.*, Journal of Educational Psychology 103:399–414, 2011; McDaniel *et al.* 2012; Roediger *et al.*, Journal of Experimental Psychology: Applied 17:382–395, 2011a). Retrieval practice, or the use of quizzes and exams to engage and enhance retrieval processes, has been widely established as an effective strategy for facilitating learning in laboratory settings (e.g., Roediger *et al.* 2011c). In this article, we review recent findings from applied research that demonstrate that retrieval practice enhances long-term classroom learning, delayed quizzes are particularly potent for retention, quizzes benefit students' transfer to novel quiz items, and quizzes with feedback improve students' learning and metacognitive awareness. In addition to generating evidence to support retrieval-based learning, these applied research studies also enhanced the professional development of the teachers, administrators, and scientists involved in the project. In this article, it is our hope that by sharing what we have learned from a variety of perspectives, applied scientific research in K-12 classrooms will continue to be explored and generated at local, state, and national levels, improving student learning and educational decision-making.

**Keywords** Applied research · Education · Teaching · Retrieval practice · Quizzing · Testing effect

What began as an innovative grant from the Institute of Education Sciences grew into a 5-year project spanning nine teachers and more than 1,400 sixth, seventh, and eighth grade

P. K. Agarwal (✉)
Department of Psychology, Washington University in St. Louis, St. Louis, MO 63130, USA
e-mail: pooja.agarwal@wustl.edu

P. M. Bain
Columbia Middle School, Columbia, IL 62236, USA

R. W. Chamberlain
College of Education, Roosevelt University, Chicago, IL 60605, USA

students from the state of Illinois. Columbia Middle School, in Columbia, IL, USA, was approached by professors from Washington University in St. Louis about forming a research partnership. When the conversations began, it was evident that this was a new endeavor for everyone—the school administrators, the teachers, the research staff, and the middle school students. Through hard work and a great deal of trial-and-error, findings from this project confirmed that retrieval practice via quizzing improves students' long-term learning in an applied setting. In addition, middle school teachers and administrators became well versed in scientifically rigorous research, while university professors and scientists gained extensive experience in the implementation of classroom-based research. In this article, it is our hope that by reviewing applied retrieval practice research and sharing what we have learned from the perspectives of a teacher, a principal, and a scientist, applied research in K-12 classrooms will continue to be explored and generated at local, state, and national levels, improving student learning and educational decision-making.

## Review of Retrieval Practice Research at Columbia Middle School

In each of the research studies conducted at Columbia Middle School, a central aim was to extend laboratory research on the retrieval practice effect (also called the testing effect) to authentic classrooms and materials, an objective rare in previous research (Roediger *et al.* 2010; Roediger *et al.* 2011b). Typically in laboratory studies, students study a set of material (e.g., word pairs, foreign language vocabulary words, prose passages), engage in retrieval practice via quizzing or testing, and then immediately or after a delay (ranging from hours, to days, to weeks), students complete a final criterial test. In general, retrieval practice improves final performance and increases retention for a variety of student populations, materials, and time delays when compared to restudying or no quizzing in laboratory research (Karpicke and Grimaldi 2012, in this issue; Roediger *et al.* 2011c).

In contrast, for applied research conducted at Columbia Middle School, students in grades 6–8 studied textbook chapter material in social studies and science (while also being presented lessons from the classroom teacher), engaged in retrieval practice via quizzes administered on a clicker response system (Ward 2007), and retention was measured after longer delays (weeks to months). The overall procedure for most studies conducted at Columbia Middle School is displayed in Fig. 1. In general, chapter material was counterbalanced across within-subject conditions (e.g., quizzed vs. not quizzed) and students completed three quizzes (a pre-quiz immediately before the teacher's lesson, a post-quiz after the teacher's lesson, and a review quiz a few days later) or simply listened to a teacher's lesson and completed classroom assignments. Due to the length of textbook chapters (e.g., Ancient China), smaller portions of the chapter were presented during lessons and quizzes; thus, short lessons and quizzes (for quizzed material) would typically continue for about 11 days, during which all chapter material would be presented and quizzed material would be quizzed a maximum of three times (see Roediger *et al.* 2011a, for additional methodological details).

To examine the potential benefits of retrieval practice on students' long-term learning, performance on quizzed material was compared to performance on non-quizzed material on final criterial exams administered at the end of the chapter (approximately 2 days after the critical manipulation was complete) and also at the end of the semester (a few months after the chapter exam, Fig. 1). As reviewed next, studies conducted at Columbia Middle School consistently demonstrated a retrieval practice effect, such that retention was better for quizzed material than for non-quizzed material. At longer delays (i.e., at the end of the

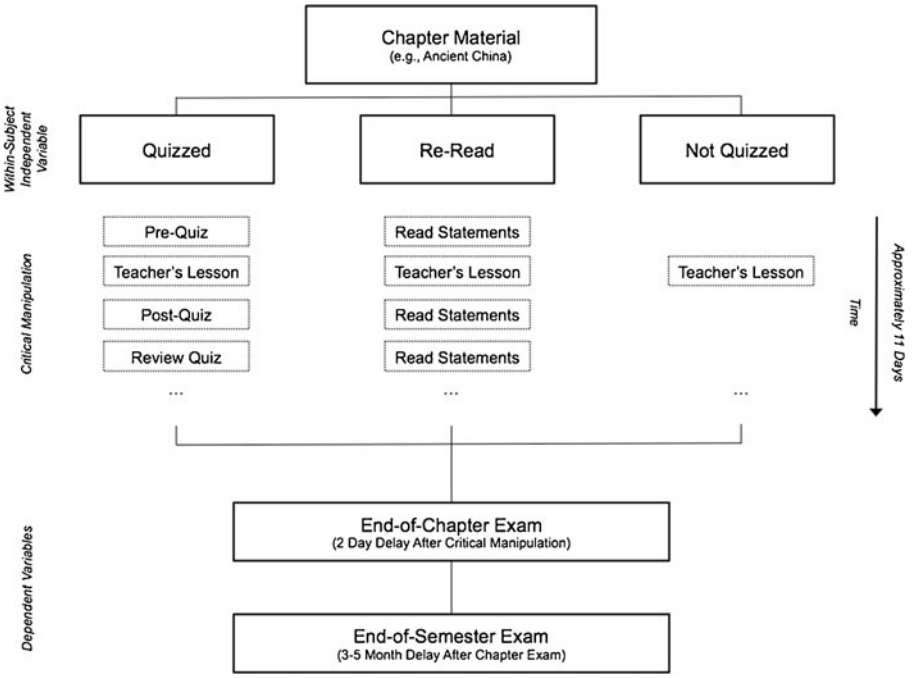


**Fig. 1** Typical experimental procedure for retrieval practice studies conducted at Columbia Middle School, adapted from methods reported in Roediger *et al.* (2011a; experiment 2)

academic semester or even school year), significant retrieval practice effects persisted, indicating long-term benefits from retrieval-based learning in an applied setting with authentic classroom materials.

## Quizzes enhance classroom learning

In one of the first studies to evaluate retrieval-based learning at Columbia Middle School (in the second author's classroom), students ($N$=142) in the sixth grade social studies received three short low-stakes quizzes (with feedback) via a clicker response system (Ward 2007; see Fig. 2). Retention was assessed at the end of social studies chapters (e.g., history of ancient Egypt, China, Africa, etc.) and also at the end of the academic semester (a longer retention interval compared to typical lab intervals; Roediger *et al.* 2011a). In the first experiment, Roediger *et al.* used a within-subject design and examined the learning of chapter material on which students received three multiple-choice quizzes compared to learning of material on which students did not receive quizzes and instead received the teacher's typical lessons and assignments (i.e., quizzed vs. not quizzed conditions; see Fig. 1). Regardless of the final chapter exam format (free recall or multiple choice), student performance was significantly greater for quizzed material than for non-quizzed material. Even at the end of the semester (approximately 1–2 months after the teacher's lessons), student performance was still significantly greater for quizzed material than non-quizzed material (79 vs. 67 %), a rigorous demonstration of the long-lasting benefits of quizzing in an applied setting (see Fig. 3).

In a second experiment (Roediger *et al.* 2011a), benefits from quizzing were compared to a re-reading control, a condition in which students' exposure time to material was equated to quizzing (as shown in Fig. 1). Again, student performance on multiple-choice chapter exams

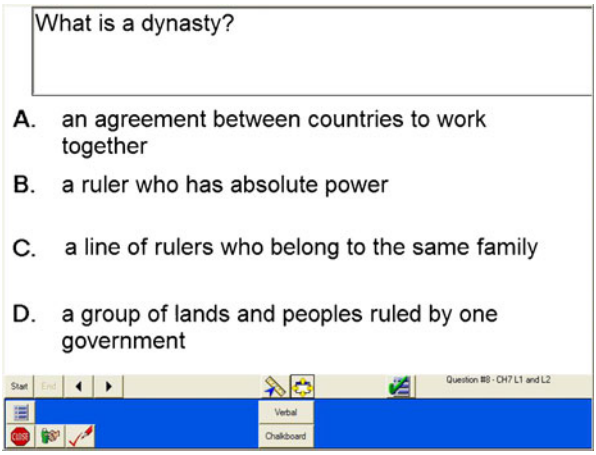


**Fig. 2** A screenshot from a 6th grade Social Studies quiz, administered via a clicker response system (Ward 2007) as part of a study by Roediger *et al.* (2011a)

was greater in the quizzed condition compared to no quizzing (91 vs. 81 %). Critically, end-of-chapter performance was also significantly greater for quizzing compared to re-reading (91 vs. 83 %), indicating that even in applied settings, the benefits of quizzing outweighed potential benefits from the re-reading of material (a common student strategy; Karpicke *et al.* 2009). As displayed in Fig. 3, these differences between conditions were no longer statistically significant at the end of the semester; however, as described by Roediger *et al.*, the differences between conditions were significant for some (but not all) of the chapter material included in the second experiment.

In a third experiment, the sixth grade social studies students ($N$=132) were able to quiz on their own at home (Roediger *et al.* 2011a). Yet again, student performance on multiple-choice chapter exams was significantly greater for quizzed material than non-quizzed

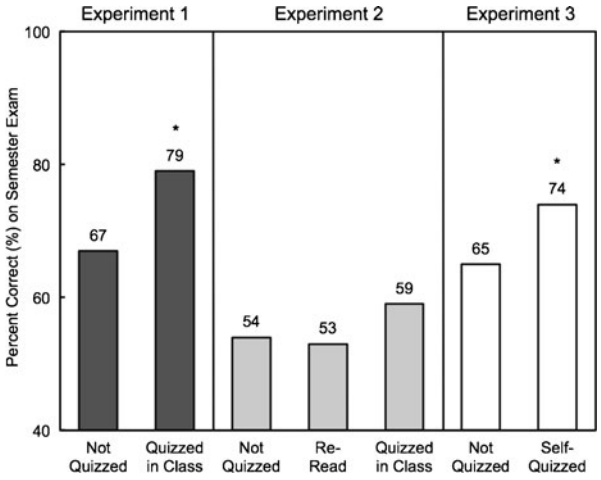


**Fig. 3** Quizzes enhance middle school students' end-of-the-semester multiple-choice exam performance, compared to re-reading or no-quiz control conditions. Figure adapted from data reported in Roediger *et al.* (2011a). *Significant, compared to the respective not quizzed condition, at an alpha level of $p<0.05$

material (90 vs. 82 %), an effect that persisted to the end of the semester (see Fig. 3). In addition, survey results from the end of the school year revealed that 97 % of students reported that the quizzes increased their learning and 65 % reported that the quizzes decreased their test anxiety, indicating that the students' themselves recognized benefits from retrieval-based learning.

This set of experiments by Roediger *et al.* clearly demonstrated the positive effects of retrieval practice on middle school students' long-term social studies performance, one of the first such findings in an authentic school setting with classroom materials. In addition, the second experiment revealed a benefit for quizzing vs. re-reading, and the third experiment demonstrated a benefit from quizzing even when students self-quizzed at home.

## Delayed quizzes improve long-term memory

A follow-up set of experiments in eighth grade science classrooms was conducted to examine the number and frequency of quizzes most beneficial for students' long-term learning (McDaniel *et al.* 2011). In this study, the first experiment ($N$=139) followed the format of the previously reviewed experiments in which students received three in-class quizzes and retention was measured on chapter and end-of-the-semester exams (Fig. 1). Importantly, retention for science content was also measured at the end of the school year, again a far longer delay (approximately 5–8 months after the teacher's lessons) than in typical lab experiments. Across the three retention intervals (i.e., end-of-chapter, end-of-semester, and end-of-year exams), student performance on quizzed material was consistently and significantly greater than performance on non-quizzed material, even at the end of the school year. Thus, McDaniel *et al.* replicated prior findings regarding the benefits of retrieval practice in authentic classrooms, but in an additional subject area (science) and with still longer retention intervals (5–8 months).

In a second set of experiments in eighth grade science (McDaniel *et al.* 2011; $N$=287 across experiments 2a and 2b), a complex within-subjects design of each of the eight possible combinations of quizzes was employed, manipulating the number and frequency with which the clicker quizzes occurred: zero, one, two, or three quizzes, crossed with quizzes occurring before the teacher's lesson (i.e., a pre-quiz), after the teacher's lesson (i.e., a post-quiz), and/or a few days after the teacher's lesson (i.e., a review quiz; see Table 1). It is worth nothing that the design for this set of experiments was accomplished in a school setting with daily student attendance, in contrast to typical lab settings where it can be difficult for college students to return after a delay. In addition, the more complex chapter material in eighth grade science (e.g., genetics, evolution, anatomy, and astronomy) allowed for completion of the critical manipulation over approximately 31 days, in contrast to about 11 days per sixth grade social studies chapter in the study by Roediger *et al.* (2011a).

As displayed in Fig. 4 (left side), chapter exam performance was significantly greater in the post-quiz and review quiz conditions, compared to the non-quizzed condition (McDaniel *et al.* 2011, experiment 2b). Chapter exam performance further increased when post- and review quizzes were combined, a significant increase from chapter exam performance following only one review quiz. At the end of the semester after a 3-month delay (Fig. 4, right side), one review quiz administered a few days after the teacher's lessons resulted in significantly greater performance compared to no quizzing. While not reported by McDaniel *et al.*, we can infer from Fig. 4 that performance in the pre-post-review condition (far right bar) would have likely been significantly greater than performance in the non-quizzed condition at the end of the semester. Contrary to notion that pre-tests enhance learning, however, a quiz administered before the teacher's lesson (i.e., the pre-quiz condition) did not

**Table 1** Eight within-subjects quiz conditions used to examine potential benefits from quiz frequency

| | | Initial quizzes | | | Chapter exam | Semester exam |
|---|---|---|---|---|---|---|
| | | $Time_1$ | $Time_2$ | $Time_3$ | | |
| 1 | Not quizzed | | | | X | X |
| 2 | Pre-quiz | X | | | X | X |
| 3 | Post-quiz | | X | | X | X |
| 4 | Review quiz | | | X | X | X |
| 5 | Pre-post | X | X | | X | X |
| 6 | Pre-review | X | | X | X | X |
| 7 | Post-review | | X | X | X | X |
| 8 | Pre-post-review | X | X | X | X | X |

Table adapted from McDaniel *et al.* (2011; experiments 2a and 2b)

$Time_1$ refers to a quiz that occurred immediately before the teacher's lesson; $time_2$ refers to a quiz that occurred immediately after the teacher's lesson, and $time_3$ refers to a quiz that occurred 24 h before the chapter exam. The end-of-semester exam occurred approximately 3 months after the chapter exam

improve performance on end-of-chapter or end-of-semester exams compared to no quizzing (cf. Karpicke and Grimaldi 2012, this issue).

Consistent with experiment 1, performance on material quizzed three times was significantly greater than performance in the not quizzed condition at the end of the school year (82 vs. 69 %), 8 months after initial learning, the largest long-term retrieval practice effect in an authentic classroom (with authentic materials) to date (McDaniel *et al* 2011, experiment 2b). Similar to survey results from the study by Roediger *et al.* (2011a), 89 % of eighth grade

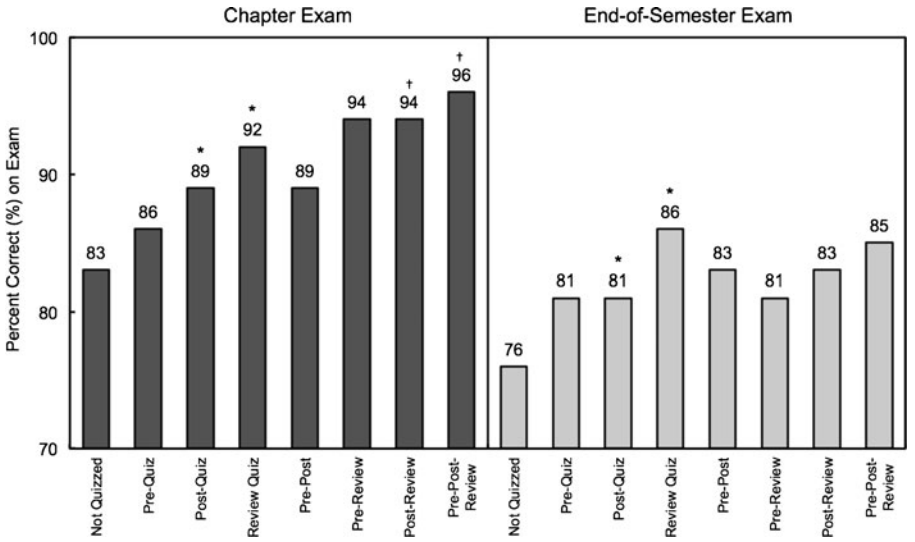


**Fig. 4** Frequent and delayed (i.e., post- and review) quizzes improve middle school students' long-term learning. Figure adapted from data reported in McDaniel *et al.* (2011, experiment 2b). *Significant, compared to the respective not quizzed condition, at an alpha level of $p<0.05$. †Significant, compared to the respective review quiz condition, at an alpha level of $p<0.05$

science students reported that the quizzes increased their learning and 64 % reported that the quizzes decreased their test anxiety before a chapter exam.

To summarize, quizzing in eighth grade science significantly improved students' long-term retention, and this retrieval practice effect was obtained with only a single review quiz administered 8 months prior to the final criterial test (McDaniel *et al.* 2011; for more information about the distribution and frequency of quizzing, see also Carpenter *et al.* 2012, and Rawson and Dunlosky 2012, both articles in this issue). In contrast to quizzing immediately before (i.e., a pre-quiz) or after (i.e., a post-quiz) a teacher's lesson, a delayed quiz (i.e., a review quiz a few days later) was most potent for enhancing long-term retention.

## Quizzes benefit students' transfer to novel quiz items

In the experiments reported thus far (McDaniel *et al.* 2011; Roediger *et al.* 2011a), quiz and exam items were identical, except multiple-choice alternatives were randomly re-ordered across quizzes and exams. In other words, a potential criticism could be that students were simply memorizing the answer to a multiple-choice item. In order to address this issue, McDaniel *et al.* (2012) conducted a set of experiments in the seventh and eighth grade science classrooms in which quiz and exam items differed. In experiment 1 with seventh grade science students ($N=142$), multiple-choice quiz questions either prompted students to provide a term given its definition (e.g., What are the smaller units that form proteins?) or to provide a definition of a concept (e.g., What are amino acids?); final criterial exam format was either consistent or different from the quizzing format (forming a $3\times2$ within-subject design, with concept quiz, definition quiz, and non-quizzed initial conditions crossed with a concept or definition chapter exam). Similarly in experiment 2 with eighth grade science students ($N=142$), multiple-choice quiz and exam questions either prompted students to provide a term given its definition (e.g., What is the term for when two or more organisms vie for limited environmental resources?) or to provide a term given a scenario or application question (e.g., A group of 500 pandas are living in a reserve. Recent dry weather has reduced the bamboo populations, which the pandas rely on. The pandas are in what type of relationship?), using a similar $3\times2$ within-subject design. In other words, items were either the same across quizzes and exams (e.g., application–application), different across quizzes and exams (e.g., concept–application), or not quizzed (e.g., not quizzed–application). The critical manipulation of chapter material (see Fig. 1) occurred over the course of 11–16 days, longer than most laboratory studies of retrieval practice.

Across the two experiments, chapter exam performance on quizzed material was significantly greater than performance on non-quizzed material (McDaniel *et al.* 2012; see Fig. 5). As displayed for experiment 1 (left side), significant retrieval practice effects were demonstrated, regardless of whether quizzes and exams were in the same format or different format. For experiment 2 (right side), however, retrieval practice effects were found for same and different formats with a concept chapter exam, but not when the chapter exam was comprised of application questions. This finding suggests that middle school students may be able to transfer knowledge from an application quiz to a conceptual exam, but not vice versa. In general, the set of applied experiments by McDaniel *et al.* demonstrates the "flexibility" or transfer of retrieval-based learning, even when quiz and final criterial items differ (see also Butler 2010).

## Quizzes with feedback improve students' learning and metacognitive awareness

Finally, Agarwal *et al.* (2009) examined the benefit of feedback on students' metacognition, or their own awareness of their learning. Many laboratory studies have demonstrated that

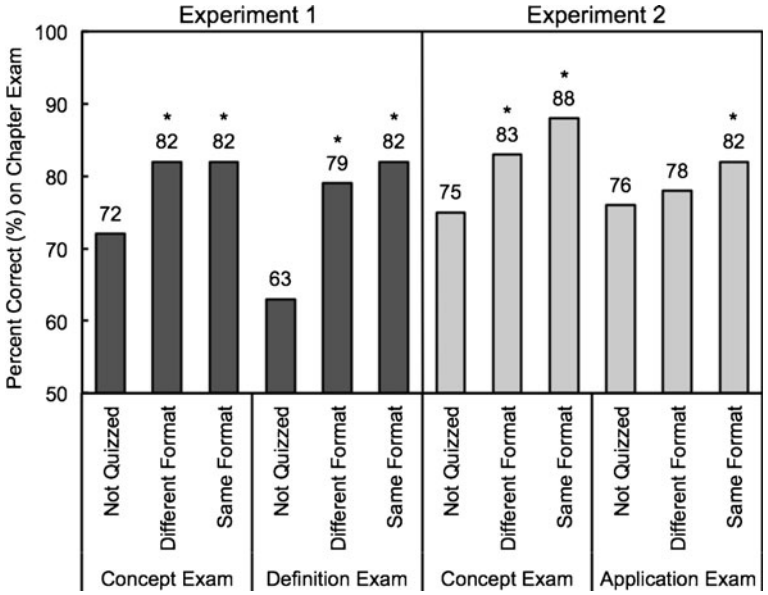


**Fig. 5** Quizzes benefit middle school students' transfer to novel (i.e., different format) quiz items. Figure adapted from data reported in McDaniel *et al.* (2012). *Significant, compared to the respective not quizzed condition, at an alpha level of $p<0.05$

feedback improves students' actual performance and also students' predictions of their performance (see Butler *et al.* 2008), and Agarwal *et al.* aimed to extend this work into an authentic school setting with classroom materials. Of particular relevance to educators, feedback is typically believed to be beneficial for only students' incorrect answers. However, in both Butler *et al.* (2008) and Agarwal *et al.* (2009), results demonstrated that feedback after quizzing improved students' *low confidence correct responses*.

In a study with seventh grade science students ($N=75$, excluding special education and absent students), students received three quizzes, except some quiz items were followed by immediate feedback and some items were *not* followed by feedback (Agarwal *et al.* 2009). A non-quizzed control condition was also included and after responding to each quiz item, students were asked to rate their confidence in their accuracy on a 1–5 scale after each quizzed item (e.g., from a guess response to a "definitely accurate" response). Surprisingly, a significant retrieval practice effect occurred on chapter exams *only* if quiz questions were followed by immediate feedback (94 vs. 80 %); chapter exam performance was similar in the quiz-without-feedback and non-quizzed conditions (83 vs. 80 %). In contrast to laboratory findings of benefits from retrieval practice without feedback (e.g., Butler *et al.* 2008), perhaps students in applied settings only reap a benefit from quizzing when followed by feedback, a speculation that warrants further research.

Further evidence of benefits from feedback were demonstrated by Agarwal *et al.* (2009), consistent with Butler *et al.* (2008), in that feedback during initial quizzes significantly improved student performance on the chapter exam for initially *correct*, but low confidence, responses (see Fig. 6). In other words, even if students were initially correct and yet reported that they were guessing, the provision of immediate feedback during quizzing increased the extent to which these low confidence correct answers were remembered on the later chapter exam. Students' resolution, or the correspondence between their quiz performance and

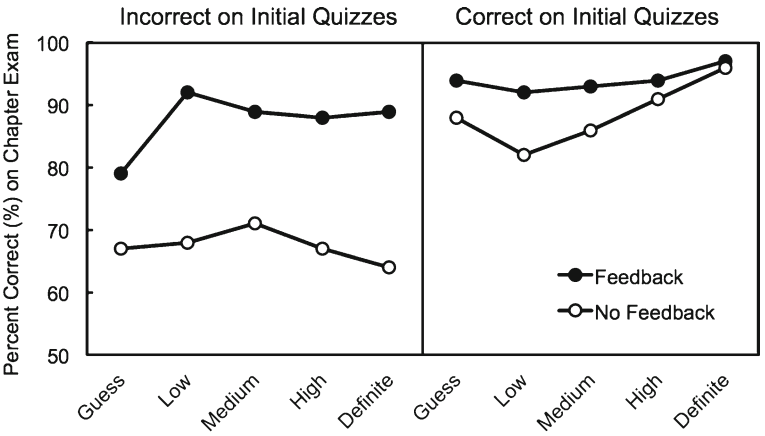


**Fig. 6** Quizzes with feedback improve middle school students' chapter exam performance on initially correct, but low confidence, responses. Figure adapted from Agarwal *et al.* (2009)

reported confidence, also improved across the three quizzes with feedback, compared to resolution across quizzes without feedback.

Applied research is essential to extending and verifying laboratory findings. Consider, however, that each of the studies reviewed took at least 4–8 months to conduct, and even longer to organize and analyze data from hundreds of students. Sometimes, data were lost and students were absent. Occasionally, snow days, fire drills, and other disruptions were unavoidable. While difficulties in applied research will continue to exist for the foreseeable future, next we describe what we (a teacher, a principal, and a scientist) have learned from conducting years of applied research at Columbia Middle School (particularly its advantages and unanticipated revelations) to encourage future implementation of applied research.

## Recommendations from a Teacher, a Principal, and a Scientist

### Thoughts from a sixth grade teacher

Research is a vital component in many disciplines (e.g., medicine, science, consumer products, advertising); yet, few teacher education courses or professional development workshops incorporate recent research on how people learn. As a veteran teacher, I (the second author) have had to personally seek out this information; it was rarely available through typical channels. When the opportunity arose that would allow me to become involved in research on how students learn, I seized the opportunity. Fortunately, the project design followed strict research procedures while also incorporating the tried and true methods of veteran teachers.

The benefits from being involved in this retrieval practice research far outweighed any drawbacks. The project allowed me the opportunity to receive a Smart Board, a clicker system, and a full-time research assistant. An additional advantage included the fact that I did not have to change my style of teaching. Most importantly, the study helped me to understand that student success was not a fluke; there was a reason *why* students in my class were successful.

What I had not anticipated was how much I would personally benefit from my involvement. My own esteem rose as I became a part of this research project and I witnessed the increased

learning and engagement of my students. I gained an awareness of educational research, the wealth of resources available to teachers, and I learned where to find this information. I was able to contribute to the writing of an Institute of Education Sciences practice guide about cognitive learning strategies (Pashler *et al.* 2007) and I have shared my experience with retrieval-based learning with fellow teachers and administrators around the country. In addition, although my participation in the research ended a few years ago, I continue to reap benefits. I became a stronger teacher both inside and outside of my classroom, and I continue to use quizzing (facilitated by technology) every day. The only drawback to participation in applied research was that I needed to plan far in advance because of all who were involved in the project and making any last-minute changes was difficult.

Involvement with a university study greatly benefits all: students, teachers, districts, and universities. The field of education is rapidly changing; teacher education degrees and curriculum directors need to be aware of the wealth of knowledge that has emerged. One can no longer fall back on the "but, that's how I was taught" mentality. Research to examine "what works" in education is essential. If we truly want to reach all students, we need to educate current and future teachers about research-based strategies that lead to authentic learning. My participation in applied research did just that—I am confident in my ability to understand, evaluate, and implement research-based strategies in my own classroom.

## Thoughts from a middle school principal

When I (the third author) was approached about Columbia Middle School's participation in a research project, my first thought was how our involvement would be beneficial for the school and most importantly for our students. I was cautious but held an open mind due to what I anticipated could be an innovative study that fit in our school's technology plan while introducing research-based strategies that could facilitate improved learning across a broad range of student abilities.

From the beginning of the project, my concern was that the research procedures might distract from classroom instruction and learning. Instead, I found that the time away from instruction for implementing quizzes and retrieval practice was well worth the investment; indeed, all students demonstrated improved performance by nearly a letter grade, including our students with individual educational plans. In addition, our school's curriculum design remained unchanged through our participation. In fact, I became *more* aware of the importance of knowing how our students were performing in order to drive instructional decision-making. Now, I am more aware of the positive impact of research on student learning, especially when research based decision-making has never been more important for principals.

Soon, principal performance will be evaluated based on student achievement; thus, it is imperative for principals to know how to analyze and understand student performance data, and change curriculum and instruction accordingly. I encourage principals to be open to research opportunities in their school. The impact our middle school's participation had on student learning was an immediate return on the investment of our time and effort throughout the project. I highly recommend that principals welcome the possibilities from applied research that promise to benefit your students, your teachers, and also your own professional development.

## Thoughts from a university scientist

As the Research Coordinator of Washington University's project at Columbia Middle School, I (the first author) became the eyes, ears, and hands of the day-to-day activities at

the middle school. Through my frequent interaction with students and teachers, coordination of 10 research assistants, development of materials, implementation of research procedures, and analysis of results, I learned quickly and avidly what is required when conducting applied research: strong relationships between the university and school settings, the ability to troubleshoot challenges regarding experimental control that present themselves in applied settings, and patience. Frequent communication was crucial to ensure that everyone's priorities and needs could be addressed and met.

My personal participation in this project has been rewarding by confirming that applied research is not only possible, but also beneficial and necessary in moving the fields of cognitive and educational psychology forward. In addition, I have learned that applied research can be accomplished at a basic scientific level. One need not compromise the merits of experimental design that allow us to establish a strong relationship between cause and effect when trying to understand a complex phenomenon such as learning. My participation also provided me with significant experience in facilitating relationships between scientists and practitioners. I was able to support communication between professors and the middle school teachers, while also anticipating the accommodations that scientists needed to make in order to maintain beneficial collaboration for all people involved. In addition, our numerous research assistants learned a great deal about applied research, with many of them continuing to pursue careers in cognitive psychology, school counseling, and curriculum development.

Finally, while serving as Director of Statewide K-12 Assessment at the Illinois State Board of Education, I was able to draw on my applied research experience when communicating assessment data to practitioners, evaluating how decisions at the state level might affect educators at the local level, and collaborating with technical advisory committees (comprised of academics) in making assessment decisions. Applied research not only benefits students, teachers, and school administrators; it also improves policy and decision-making at a state level.

## Recommendation

The three of us agree: applied research is valuable for all those involved—the teachers, the principals, the scientists, and first and foremost, the students. Originally, we did not anticipate the impact that teachers' and the principal's participation in research would have on their professional growth: they made educational decisions based on the research findings and it was not uncommon to see them discover resources beyond the project that brought additional benefit to students. In addition, the teachers and administrators asked to share what they discovered with their peers, a gratifying continuation of the project.

From a research perspective, we were able to bridge the gap between lab research and the "real world" of education with one of the first projects in the field of cognitive psychology using actual classroom materials, curricula, schedules, students, and teachers. Even with the everyday adjustments required in an applied setting, results demonstrating the effectiveness of retrieval practice were reliable, publishable, and informative for future research.

In sum, while there are advantages and drawbacks to applied research, we hope that teachers, principals, scientists, policy makers, and others will consider getting involved. We are confident that the value and experience of school research is one we would not have had if we all originally allowed our concerns to distract from this growing area of research. In the end, we have been able to significantly improve student learning and achievement through retrieval-based learning—an end goal around which we all can rally.

**Acknowledgments** This research was supported by the Institute of Education Sciences, U.S. Department of Education through grants R305H060080 and R305A110550 awarded to Washington University in St. Louis. The opinions expressed are those of the authors and do not represent the views of the Institute or the US Department of Education.

We thank our colleagues Henry L. Roediger, III, Mark A. McDaniel, and Kathleen B. McDermott for their contributions to the research reviewed, as well as participating students, teachers, and parents from Columbia, IL, USA. We also thank an anonymous reviewer for helpful comments on an earlier version of this manuscript.

## References

Agarwal, P.K., Roediger, H.L., McDaniel, M.A., McDermott, K.B. (2009, November). *Feedback increases middle school students' resolution and retention of correct responses*. Poster presented at the 50th Annual Meeting of the Psychonomic Society, Boston, MA.

Butler, A. C. (2010). Repeated testing produces superior transfer of learning relative to repeated studying. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 1118–1133.

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

Carpenter, S.K., Cepeda, N.J., Rohrer, D., Kang, S.H.K., Pashler, H. (2012). Using spacing to enhance diverse forms of learning: Review of recent research and implications for instruction. *Educational Psychology Review.* doi:10.1007/s10648-012-9205-z.

Karpicke, J. D., Grimaldi, P. J. (2012). Retrieval-based learning: A perspective for enhancing meaningful learning. *Educational Psychology Review.* doi:10.1007/s10648-012-9202-2.

Karpicke, J. D., Butler, A. C., & Roediger, H. L. (2009). Metacognitive strategies in student learning: Do students practice retrieval when they study on their own? *Memory, 17*, 471–479.

McDaniel, M. A., Agarwal, P. K., Huelser, B. J., McDermott, K. B., & Roediger, H. L. (2011). Test-enhanced learning in a middle school science classroom: The effects of quiz frequency and placement. *Journal of Educational Psychology, 103*, 399–414.

McDaniel, M.A., Thomas, R.C., Agarwal, P.K., McDermott, K.B., Roediger, H.L. (2012). *Quizzing promotes transfer of target principles in middle school science: Benefits on classroom exams*. Manuscript submitted for publication.

Pashler, H., Bain, P.M., Bottge, B.A., Graesser, A., Koedinger, K., McDaniel, M.A., *et al.* (2007). *Organizing instruction and study to improve student learning* (NCER 2007-2004). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education. Available online at http://ies.ed.gov/ncee/wwc/pdf/practice_guides/20072004.pdf.

Rawson, K. A., & Dunlosky, J. (2012). When is practice testing most effective for improving durability and efficiency of student learning? *Educational Psychology Review.* doi:10.1007/s10648-012-9203-1.

Roediger, H. L., Agarwal, P. K., Kang, S. H. K., & Marsh, E. J. (2010). Benefits of testing memory: Best practices and boundary conditions. In G. M. Davies & D. B. Wright (Eds.), *New frontiers in applied memory* (pp. 13–49). Brighton: Psychology Press.

Roediger, H. L., Agarwal, P. K., McDaniel, M. A., & McDermott, K. B. (2011a). Test-enhanced learning in the classroom: Long-term improvements from quizzing. *Journal of Experimental Psychology: Applied, 17*, 382–395.

Roediger, H. L., McDermott, K. B., & McDaniel, M. A. (2011b). Using testing to improve learning and memory. In M. A. Gernsbacher, R. Pew, L. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 65–74). New York: Worth Publishing Co.

Roediger, H. L., Putnam, A. L., & Smith, M. A. (2011c). Ten benefits of testing and their applications to educational practice. In J. Mestre & B. Ross (Eds.), *Psychology of learning and motivation: Cognition in education* (pp. 1–36). Oxford: Elsevier.

Ward, D. (2007). *eInstruction: Classroom performance system [computer software]*. Denton: eInstruction Corporation.