Annales Mathematicae et Informaticae
Accepted manuscript
DOI: 10.33039/ami.2024.05.002
URL: https://ami.uni-eszterhazy.hu

Retrieval practice – a tool to be able to retain higher mathematics even 3 months after the exam*

Anna Muzsnay[ae], **Csaba Szabó**[be], **Janka Szeibert**[cd]

[a]University of Debrecen
muzsnay.anna@science.unideb.hu

[b]Eötvös Loránd University, Faculty of Science
szabo.csaba.mathdid@ttk.elte.hu

[c]Eötvös Loránd University, Faculty of Primary and Pre-School Education
szeibert.janka@gmail.com

[d]Edutus University

[e]MTA-ELTE Theory of Learning Mathematics Research Group

Abstract. It is a common phenomenon that students forget the learned material within a few days after their exam. A considerable part of university students do not gain long-term knowledge. Aiming to reduce forgetting and increase further retention in a first-year mathematics course for mathematics pre-service teachers, we applied a special kind of retrieval practice in their lessons. The positive effects of retrieval practice – the strategic use of retrieval to enhance memory – have been shown in the medium term in learning university mathematics. In this paper, we investigate the potential benefit of the applied retrieval practice in learning Number Theory at the university level, focusing on knowledge lasting for 3 months. $N = 42$ first-year pre-service mathematics teacher students wrote a post-test on the material they learned in the course Number Theory three months after their exam. According to our results, those, who learned Number Theory by retrieval practice, performed significantly better than those who learned on the traditional way. Our findings suggest that retrieval practice can have a powerful, long-lasting effect on learning and solving complex mathematical problems.

*This research was supported by the Research Programme for Public Education Development of the Hungarian Academy of Sciences and by Digital Education Development Competence Center at Eötvös Loránd University, Budapest, with project number 2022-1.1.1-KK-2022-00003.

Submitted: December 8, 2023
Accepted: May 8, 2024
Published online: May 10, 2024

Keywords: retrieval practice, test-enhanced learning, active recall, university mathematics, Number Theory

1. Introduction

Many studies showed that a significant part of university students chose learning strategies that are not beneficial in creating long-term knowledge [8]. Without applying powerful learning techniques, students forget the learned material within a few days after their exam. However, gaining long-term knowledge is crucial in learning and teaching mathematics [11]. This is especially true for future mathematics teachers since they are the ones who will teach future students [18]. A possible tool to increase further retention is retrieval practice, the strategic use of retrieval to enhance memory. In this study, we aim to show that by using a simple intervention – implementing a special kind of retrieval practice – educators can support their incoming students and facilitate the creation and retention of knowledge and problem-solving skills in a given higher mathematical content.

2. Literature review

Retrieval practice – the strategic use of retrieval to enhance memory – has been proven to be an effective learning method in many cases [7, 29]. Retrieval practice, also known as testing or test-enhanced learning, can refer to any activity (such as questions during class, quizzes, flashcards, brain dump, and examination questions) that requires retrieving information from memory without the help of any external sources. In the past two decades, numerous studies have shown that by actively recalling information from memory one can get more durable knowledge than by rereading the material [28]. Information learned by practicing recalling is retained significantly better than information learned by non-retrieval-based strategies, such as copying [4], re-reading [24, 28], or organizing the information in a new way [10].

Although retrieval practice is not a new concept – it was first studied more than a hundred years ago (for early research see [1]) – it has received more attention in psychological and educational research only in the last 20 years. Test-enhanced learning has been proven to be a successful learning strategy in many different areas as it contributes to consolidating the constructed mental representations in memory (e.g., [14]). The testing effect – the phenomenon that retrieving information from short- or long-term memory can strengthen one's memory of the retrieved information – has been shown in many different areas such as memorizing texts, foreign language vocabulary, general knowledge facts, learning materials that include visual or spatial information, and also in skill learning [6, 26, 28]. It was also demonstrated in laboratory circumstances and real educational environments [7, 21, 25]. Even though the research on retrieval practice indicates that it is a powerful way to promote learning, some researchers have suggested that the success of retrieval practice depends on the complexity of the to-be-learned material.

2.1. Retrieval practice and the complexity of the learning material

It is not evident if the testing effect is present when learning "complex" materials. On the one hand, retrieval works against forgetting; knowledge acquired by retrieval practice leads to a lower forgetting rate [15, 23, 31]. On the other hand, since students rarely remember all content items of the to-be-learned material perfectly (e.g., [28]) it can reduce the amount of successfully executed knowledge construction activities [27]. In their study, Roelle and Berthold examined whether the effects of incorporating retrieval into learning tasks depend on the learning tasks' complexity. They argue that the benefit of incorporating retrieval into learning tasks depends on the complexity of the tasks – the degree to which learning tasks require learners to combine content items that are included in the learning material ([27, p. 142]). They found that the net benefit of incorporating retrieval was higher for the low-complexity tasks, and as the complexity of the learning task increases the benefit decreases.

Also, Van Gog and Sweller claim that retrieval practice is not beneficial for complex materials [9]. According to them complex materials, the testing effect either diminishes or completely disappears. In their study "complex material" is referred to as "high in element interactivity, containing various information elements that are related and must therefore be processed simultaneously in working memory" ([9, p. 248]). However, the findings of van Gog and Sweller were criticized by Karpicke et al. [13] for the lack of an objective measure of complexity. Also, they listed a series of studies where the retrieval effect was present when learning complex materials such as the research of McDaniel et al. [20], Chan [5], and Butler [3].

2.2. Retrieval practice and mathematics

In recent years there has been a growing interest in investigating the testing effect in mathematics learning, in mathematical problem-solving. Since mathematical problems require developed deductive and problem-solving skills, and the problems themselves are quite complex, it is not obvious whether or not incorporating retrieval practice is a powerful way to enhance learning in this field. Developing problem-solving skills in mathematics requires not only the memorization of facts and procedures but also a deep conceptual understanding. Although more applied research is needed in this field [2], recent studies suggest that increasing retrieval practice may be an effective way of learning mathematics [12, 16, 17, 19, 30, 32]. The experiments of Yeo and Fazio [32] and Lyle et al. [16] are particularly relevant for this study.

The study of Yeo and Fazio [32] investigated the effect of retrieval practice versus (re)studying worked examples in mathematical problem-solving five minutes and one week after a one-session learning phase. They found that the optimal learning strategy depends on the learning objectives, the retention interval, and the kind of knowledge being learned (stable facts or flexible procedures) as well.

Among other things, they showed that when the goal was to learn a novel math procedure, the effectiveness of the two methods depended on the retention interval. When they tested participants' knowledge five minutes after the learning phase using nonidentical learning problems in the test, repeated studying was more effective than repeated testing. However, one week later, the group that learned by repeated testing performed as well as the group that learned by (re)studying worked examples. With identical learning problems repeated testing was more advantageous than repeated studying.

Lyle et al. [16] investigated the effect of retrieval practice in a genuine educational setting. They measured the impact of spaced versus massed retrieval and their impact on long-term retention in a precalculus course for engineering students. They found that increasing the spacing of practice – even though it significantly reduced quiz performance – resulted in better retention at the end of a precalculus course and also 1 month later.

3. Research focus and research question

In a former study, the implementation of retrieval practice in university mathematics was investigated [30]. Their results showed that in a first-year Number-Theory course, students who learned by retrieval performed better on the final test. This result strengthens the results of Lyle et al. [16] in their precalculus course in the sense that the exam included a major part of the material learned at the beginning of the semester. In the present study, we concentrate on pre-service mathematics teachers' knowledge in a given Number Theory material three moths after they took the Number Theory exam. Our research question is the following:

Can the retrieval effect be demonstrated in a longer term, a few months after finishing the course in university mathematics (Number Theory)?

4. Methods

4.1. Sample and study design

The authors conducted a quasi-experimental study to figure out whether applying a certain form of retrieval practice leads to better knowledge in the long term in a first-year mathematics course (Algebra and Number Theory 1.). Participants of the study were first-year mathematics pre-service teacher students from University Name who took the compulsory course Algebra and Number Theory 1. in their first semester. Altogether 114 students attended the course. Among the 114 students, 46 dropped out, and 68 wrote the post-test. Their ages were between 18 and 23.

The course consisted of a 60-minute lecture and a 90-minute problem session and lasted 13 weeks. The students attended the same lectures, while their practice sessions were taught in groups of 15 students on average. All together there were six practice groups. Students learned the same material and solved the same problems during practice sessions. Three of the six groups were randomly selected as the experimental group, and the other three were the control group. We tried to eliminate the teacher's effect as much as possible: the teachers in the control and experimental groups were matched in the sense that there was one experienced teacher, one demonstrator, and one doctoral student in the control and experimental groups. Furthermore, teachers of the practice groups had a short meeting each week where they discussed the main issues related to the course. The structure of the problem-solving sessions was similar in each group.

The control group started the practice sessions by writing a short test on the material from the previous week's lecture (as it is traditional in the case of this subject). After the short test, they discussed the homework which was followed by the main part of the session: the problem-solving part with the aid of the professors.

The experimental group's sessions started by discussing the homework, then they had the problem-solving part which was followed by an end-of-class test on the material of the given practice session. So the structure of the lesson was almost identical, the difference between the two types of groups was that in the experimental group, there was no test at the beginning of the lesson, instead, they had a test at the end of the class. The end-of-class tests consisted of two open-ended problems, similar to those encountered during the practice session. We tried to ask desirably difficult questions for the students: neither too easy nor too hard for them. Students had to solve it on their own, without any help. This way, they had to retrieve what they just learned. They did not get any feedback about the solution to the problems asked in the end-of-class test, only if they explicitly asked for it.

In both groups, the tests were evaluated, and students could gain 2 points on each test. To make sure that students took these tests seriously, they needed to score 50 % by the end of the semester to pass the course.

4.2. The material covered by the course

The regular course materials for the Algebra and Number Theory lectures and problem-solving seminars were used based on the textbook by [22]. Topics covered by the course were:

- Divisibility, the greatest common divisor, the Euclidean algorithm, prime numbers, and the fundamental theorem of arithmetic.
- Special arithmetic functions, additive and multiplicative arithmetic functions. Divisor sum of multiplicative functions. The Möbius function. Perfect numbers.
- Congruences. The Euler-Fermat theorem. Linear congruences and diophantine equations. Linear congruence systems. Applications in computational number theory.

- Congruences of higher degree. Reduction to prime power, resp. prime moduli. The number of solutions, the reduction of degree in case of prime moduli. Wilson's theorem. Binomial congruences, order, primitive roots, index.

4.3. Instruments

Students' knowledge was measured at the beginning of their studies, and three months after their final exam. Before the course started, each student completed a competence-level test, which served as an input test. This test is obligatory for every pre-service mathematics teacher student and serves as a general competence and knowledge test on the high-school curriculum. Also, students who passed the course wrote a "surprise" test three months after the final exam. The test consisted of four problems and 20 minutes were given to complete it. Two problems, Problem 1 and Problem 2 involve tasks that can be solved procedurally, Problems 3 and Problem 4 are more complex. Here, we present the four problems.

Problem 1. Find the remainder of $2346235^{226688442}$ mod 23.

In this problem the only necessary knowledge is the Euler-Fermat theorem. In the solution you need to take the base mod 23 and the exponent mod $\varphi(23)$. As 23 is a prime, $\varphi(23) = 22$. At first sight, it seems slow to find the remainders. Since the test test took only 20 minutes, there was no time to use the division algorithm. But, if you look carefully at the actual numbers, pairing the digits there are numbers divisible by 23 and 22, respectively, so it can be done fast.

Problem 2. Solve the following system of congruences: $2x \equiv 8 \bmod 14$ and $x \equiv 7 \bmod 11$.

The standard solution of this problem is not very fast either. The first step is to reduce $2x \equiv 8 \bmod 14$ to $x \equiv 4 \bmod 7$, and then observe that 18 is a solution. Then, referring to the Chinese remainder theorem one can see, that 18 is the unique solution mod 77.

Problem 3. Find all solutions of the following equation over the integers:

$$3x^{16} - 4y^{48} + 17z^{2012} = 34172.$$

This problem is a challenging one. There are 4 terms, one of them is constant, and students have an arsenal of tricks for handling high-degree Diophantine equations. Here, the find a prime p and take the equation mod p trick works. The question is, which prime? This problem is a routine problem when we learn the trick, but rather tricky in this environment.

Problem 4. For which a, b digits can $\overline{a97531ba}$ be divided by 55?

This problem requires to know the divisibility rules mod 5 and 11. Since it is divisible by 5, $a = 0$ or $a = 5$. Then by taking the alternating sum of the digits b can be given. Although this is high-school knowledge, surprisingly, quite a few people got low scores on this problem.

5. Results

During the statistical analysis, we analyzed data from students who attended Algebra and Number Theory 1. and wrote the input and output tests. Altogether 79 passed the course, 68 of them took the post test. Out of the 68 students 51 wrote the pre test. The remaining 17 students can be divided into two groups. The first group took Number Theory 1. second time, they did not complete the course in their earlier studies. The second group did not write the pre-test, and this way they had to take a bridging course. We excluded the data of 4 students who had entered the university at least two years before the course. We have also excluded those 3 students who scored 0. They did not take seriously the test and handed in empty sheets. We excluded 4 excellent students who scored above 90 % on both tests and also on the midterms during the semester. Considering their test scores in the statistics would have given a false result. They scored well independently of the testing effect. In this study, we analyze the remaining 42 students' test results: 21 from the experimental group and 21 from the control group. The data analysis was conducted using R 4.2.3 software.

When analyzing the results of the input test we found that the variances of the two groups differed ($\mathrm{M_{control}} = 69.2$, $\mathrm{SD_{control}} = 11.0$, $\mathrm{M_{experimental}} = 58.9$, $\mathrm{SD_{experimental}} = 21.1$), so we applied the Welch's t-test. The test showed that there was no difference between the two groups, Welch two-way sample $t(31.5) = 1.99$, $p = 0.11$ at the beginning of the course, with effect size $d = 0.60$.

Then, we analyzed the results of the post-test. The variances were different by the F-test ($p = 0.008$), so we applied the Welch test. There was a significant difference between the two groups, $t(43.4) = 3.15$, $p < 0.001$, with effect size $d = 0.88$. The average of the experimental group was 75 % and the average of the control group was 60 % (see Figure 1).

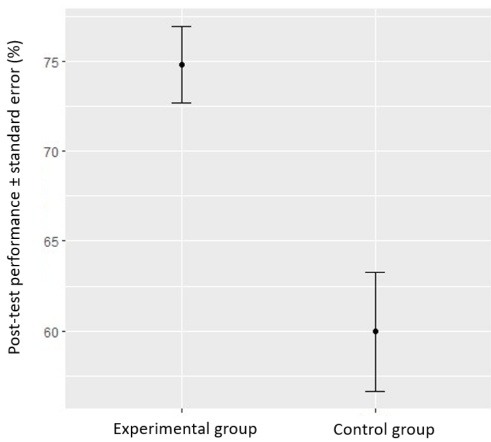

Figure 1. The results of the post-test in the experimental and the control groups.

6. Discussion

Many studies showed that a significant part of university students chose learning strategies that are not beneficial in creating long-term knowledge [8]. Without applying powerful learning techniques, students forget the learned material within a few days after their exam. However, gaining long-term knowledge is crucial in learning and teaching mathematics [11]. Retrieval practice, the strategic use of retrieval to enhance memory is recommended as an effective method for improving learning [6, 7]. It is not obvious whether the retrieval effect is present when learning "complex" materials. Furthermore, the long-term effect of retrieval practice when learning higher mathematics is still unknown.

In this case study, we investigated the effectiveness of a particular type of retrieval practice in a university mathematics course concentrating on preserving the konwledge for three months. Participants of the study were 44 first-year mathematics pre-service teacher students who took the compulsory course Algebra and Number Theory 1. Students' input level was measured by a competence-level test at the beginning of the course. The effects of retrieval practice and traditional learning were measured by a "surprise" post-test on the material they learned in the course three months after their final exam. According to our results, those, who learned Number Theory by retrieval practice, performed significantly better than those who learned on the traditional way. Our findings suggest that retrieval practice can have a powerful, fairly long-lasting effect on learning and solving complex mathematical problems.

We believe that the strength of this particular study is that we could measure the effect of a learning method in a real educational environment lasting for a whole semester. Tracking university students' knowledge can be challenging in a real school environment since reaching students after we no longer teach them is difficult and nearly impossible.

Although according to the results of this case study, this type of retrieval practice seems to be an effective way to create rather long-lasting knowledge in university mathematics, further research is needed in this area to draw far-reaching conclusions. It would be important to test the method in different school settings: with different students, in different universities, and other mathematics courses. Finally, it would be interesting to explore more deeply the effect of retrieval practice on higher mathematical knowledge in an even longer term.

References

[1] E. E. Abbott: *On the analysis of the factor of recall in the learning process*, The Psychological Review: Monograph Supplements 11.3 (1909), pp. 159–177, doi: 10.1037/h0093018.

[2] P. Agarwal, L. Nunes, J. Blunt: *Retrieval Practice Consistently Benefits Student Learning: a Systematic Review of Applied Research in Schools and Classrooms*, Educational Psychology Review 33 (Dec. 2021), pp. 1–45, doi: 10.1007/s10648-021-09595-9.

[3] A. Butler: *Repeated Testing Produces Superior Transfer of Learning Relative to Repeated Studying*, Journal of experimental psychology. Learning, memory, and cognition 36 (Sept. 2010), pp. 1118–33, doi: 10.1037/a0019902.

[4] S. Carpenter, T. Lund, C. Coffman, P. Armstrong, M. Lamm, R. Reason: *A Classroom Study on the Relationship Between Student Achievement and Retrieval-Enhanced Learning*, Educational Psychology Review 28 (May 2015), doi: 10.1007/s10648-015-9311-9.

[5] J. C. K. Chan: *Long-term effects of testing on the recall of nontested materials*, Memory 18.1 (2010), pp. 49–57, doi: 10.1080/09658210903405737.

[6] G. M. Donoghue, J. A. C. Hattie: *A Meta-Analysis of Ten Learning Techniques*, Frontiers in Education 6 (2021), issn: 2504-284X, doi: 10.3389/feduc.2021.581216.

[7] J. Dunlosky, K. Rawson, E. Marsh, M. Nathan, D. Willingham: *Improving Students' Learning With Effective Learning Techniques*, Psychological Science in the Public Interest 14 (Jan. 2013), pp. 4–58, doi: 10.1177/1529100612453266.

[8] J. Dunlosky, K. A. Rawson, E. J. Marsh, M. J. Nathan, D. T. Willingham: *What Works, What Doesn't*, Scientific American Mind 24 (2013), pp. 46–53.

[9] T. van Gog, J. Sweller: *Not New, but Nearly Forgotten: the Testing Effect Decreases or even Disappears as the Complexity of Learning Materials Increases*, Educational Psychology Review 27.2 (2015), pp. 247–264, issn: 1040726X, 1573336X, (visited on 12/04/2023).

[10] N. Goossens, G. Camp, P. Verkoeijen, H. Tabbers, S. Bouwmeester, R. Zwaan: *Distributed Practice and Retrieval Practice in Primary School Vocabulary Learning: A Multi-classroom Study: Distributed Practice and Retrieval Practice*, Applied Cognitive Psychology 30 (Oct. 2016), doi: 10.1002/acp.3245.

[11] R. F. Hopkins, K. B. Lyle, J. L. Hieb, P. A. S. Ralston: *Spaced Retrieval Practice Increases College Students' Short- and Long-Term Retention of Mathematics Knowledge*, Educational Psychology Review 28.4 (2016), pp. 853–873, (visited on 12/07/2023).

[12] R. F. Hopkins, K. B. Lyle, J. L. Hieb, P. A. S. Ralston: *Spaced Retrieval Practice Increases College Students' Short- and Long-Term Retention of Mathematics Knowledge*, Educational Psychology Review 28.4 (2016), pp. 853–873.

[13] J. D. Karpicke, W. R. Aue: *The Testing Effect Is Alive and Well with Complex Materials*, Educational Psychology Review 27 (2015), pp. 317–326, doi: 10.1007/s10648-015-9309-3.

[14] J. D. Karpicke: *2.27 - Retrieval-Based Learning: A Decade of Progress*, in: Learning and Memory: A Comprehensive Reference (Second Edition), ed. by J. H. Byrne, Second Edition, Oxford: Academic Press, 2017, pp. 487–514, isbn: 978-0-12-805291-4, doi: 10.1016/B978-0-12-809324-5.21055-9, url: https://www.sciencedirect.com/science/article/pii/B9780128093245210559.

[15] O. Kliegl, K.-H. Bäuml: *Retrieval Practice Can Insulate Items Against Intralist Interference: Evidence From the List-Length Effect, Output Interference, and Retrieval-Induced Forgetting*, Journal of Experimental Psychology: Learning, Memory, and Cognition 42 (Aug. 2015), pp. 202–214, doi: 10.1037/xlm0000172.

[16] K. Lyle, C. Bego, R. Hopkins, J. Hieb, P. Ralston: *How the Amount and Spacing of Retrieval Practice Affect the Short- and Long-Term Retention of Mathematics Knowledge*, Educational Psychology Review 32 (Mar. 2020), pp. 277–295, doi: 10.1007/s10648-019-09489-x.

[17] K. B. Lyle, N. A. Crawford: *Retrieving Essential Material at the End of Lectures Improves Performance on Statistics Exams*, Teaching of Psychology 38.2 (2011), pp. 94–97, doi: 10.1177/0098628311401587.

[18] L. Ma: *Knowing and Teaching Elementary Mathematics: Teachers' Understanding of Fundamental Mathematics in China and the United States*, Jan. 2020, isbn: 9781003009443, doi: 10.4324/9781003009443.

[19] B. M. May: *Effects of spaced, repeated retrieval practice and test-potentiated learning on mathematical knowledge and reasoning*, International Journal of Mathematical Education in Science and Technology 53.1 (2022), pp. 92–107, doi: 10.1080/0020739X.2021.1961034.

[20] M. Mcdaniel, D. Howard, G. Einstein: *The Read-Recite-Review Study Strategy*, Psychological science 20 (Apr. 2009), pp. 516–22, doi: 10.1111/j.1467-9280.2009.02325.x.

[21] K. McDermott, P. Agarwal, L. D'Antonio, H. Roediger, M. Mcdaniel: *Both Multiple-Choice and Short-Answer Quizzes Enhance Later Exam Performance in Middle and High School Classes*, Journal of Experimental Psychology Applied 20 (Nov. 2013), doi: 10.1037/xap0000004.

[22] I. Niven, H. S. Zuckerman, H. L. Montgomery: *An Introduction to the Theory of Numbers, 5th edition*, Wiley, 1991, isbn: 0471625469.

[23] M. Racsmány, A. Keresztes: *Initial Retrieval Shields Against Retrieval-Induced Forgetting*, Frontiers in psychology 6 (May 2015), p. 657, doi: 10.3389/fpsyg.2015.00657.

[24] H. Roediger, P. Agarwal, M. Mcdaniel, K. McDermott: *Test-Enhanced Learning in the Classroom: Long-Term Improvements From Quizzing*, Journal of experimental psychology. Applied 17 (Nov. 2011), pp. 382–95, doi: 10.1037/a0026252.

[25] H. Roediger, A. Putnam, M. Sumeracki: *Ten Benefits of Testing and Their Applications to Educational Practice*, in: vol. 55, Jan. 2011, pp. 1–36, isbn: 9780123876911, doi: 10.1016/B978-0-12-387691-1.00001-6.

[26] H. L. Roediger, A. C. Butler: *The critical role of retrieval practice in long-term retention*, Trends in Cognitive Sciences 15.1 (2011), pp. 20–27, issn: 1364-6613, doi: 10.1016/j.tics.2010.09.003.

[27] J. Roelle, K. Berthold: *Effects of Incorporating Retrieval Into Learning Tasks: The Complexity of the Tasks Matters*, Learning and Instruction 49 (Jan. 2017), pp. 142–156, doi: 10.1016/j.learninstruc.2017.01.008.

[28] C. Rowland: *The Effect of Testing Versus Restudy on Retention: A Meta-Analytic Review of the Testing Effect*, Psychological bulletin 140 (Aug. 2014), doi: 10.1037/a0037559.

[29] A. Saidat, P. Baker: *Effects of Worked Example on Students' Learning Outcomes in Complex Algebraic Problems*, International Journal of Instruction 16 (Apr. 2023), pp. 229–246, doi: 10.29333/iji.2023.16214a.

[30] C. Szabo, C. G. Zámbó, A. Muzsnay, J. Szeibert, L. Bernáth: *Investigación de la eficacia de práctica de recuperación en matemáticas universitarias*, Revista de Educación Superior 401 (July 2023), pp. 79–96, doi: 10.4438/1988-592X-RE-2023-401-584.

[31] K. Szpunar, K. McDermott, H. Roediger: *Testing During Study Insulates Against the Buildup of Proactive Interference*, Journal of experimental psychology. Learning, memory, and cognition 34 (Nov. 2008), pp. 1392–9, doi: 10.1037/a0013082.

[32] D. J. Yeo, L. K. Fazio: *The Optimal Learning Strategy Depends on Learning Goals and Processes: Retrieval Practice Versus Worked Examples*, Journal of Educational Psychology 111 (2019), pp. 73–90.