

Retrieval practice of a hierarchical principle structure in university introductory physics: Making stronger students

Vegard Gjerde¹, Bodil Holst, and Stein Dankert Kolstø¹

Department of Physics and Technology, University of Bergen, 5007 Bergen, Norway

 (Received 22 November 2019; accepted 10 March 2020; published 21 April 2020)

Introductory physics is taught to several hundred thousand university students every year. It is seen as especially difficult by many and the failure rate is often high. A relevant question is whether one can increase the success rate among the weaker students? Retrieval practice is an established learning strategy with large benefits. However, as pointed out last year in this journal, hardly any systematic research has been done on retrieval practice in physics. Here we present a novel tool for retrieval practice in physics called the hierarchical principle structure for mechanics (HPSM). HPSM hierarchically organizes the essential principles, equations, and definitions for translational, rotational, and fluid mechanics, to emphasize meaningful connections. We investigated HPSM in a two-phase study. First, we present a randomized controlled experiment showing that 70 min of retrieval practice of HPSM had a very large effect on a declarative factual test compared to 70 min of problem study, $d = 1.42$. In the second phase, which was carried out the following year, we implemented distributed retrieval practice of HPSM in the first 15 min of 16 lectures. Although difficult to disentangle the effect from the lectures it was embedded in, distributed retrieval practice of HPSM seems to promote factual knowledge ($r = 0.44$) and better exam results for the weaker students (significant main and interaction effects).

DOI: [10.1103/PhysRevPhysEducRes.16.013103](https://doi.org/10.1103/PhysRevPhysEducRes.16.013103)

I. INTRODUCTION

Some learning strategies are more effective than others [1], and retrieval practice is one of the learning strategies with the most positive evidence [1–11]. The effects from retrieval practice are so robust across different contexts, both in labs and in applied settings, that many cognitive scientists recommends its use for education [8,12]. However, there is still a need for more research on retrieval practice in educational settings [1,12], and especially in physics education where it has hardly been studied at all as pointed out in a recent publication in this journal [11]. Research in education, including physics education, still focuses more on encoding processes than on retrieval [13]. There seems to be an underlying fear that students will acquire disconnected facts, and ultimately have lower understanding [14]. Even Dunlosky *et al.*, who strongly encourage efforts to improve memory for facts, supply caveats against “robotically memorizing facts” [1]. In a rare case of research on retrieval practice in physics, an advantage was found for retrieval practice over peer discussion of conceptual instruction [11]. In this study,

the students watched a video lecture about speed and energy conservation, and then did either retrieval-based or peer discussion-based restudy of the content in the lecture. In a more recent study, Gjerde *et al.* found positive effects on problem solving performance from having engaged in retrieval practice of physics principles [15]. We did not find any published systematic study of retrieval practice for memorizing essential principles and definitions in physics, even though presumably this is more effective than having no specific method.

The probability of being able to remember a fact is dependent on the activation of the fact. Activation is an additive function of the base strength of the memory and associative activation from contextual cues [16,17]. When first encoded, a new fact has low base strength and few associative ties to contextual cues. The act of retrieving a memory increases the base strength of the memory [18,19]. This memory strength is a function of the recency and frequency of practice, and reflects past usefulness [16,20]; the more a fact is retrieved, the more likely it will be useful in the future. The increased strength of memories due to retrieval practice may make memories less context dependent for future recall, promote insight, inference, and generalization [21], and may enable students to use the practiced information more flexibly when meeting new concepts thereby potentiating further learning [5,22,23]. In support of these claims, retrieval practice seems to be better than restudy for tests of transfer of knowledge [4,22,24,25].

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Further, the base strength of memory chunks seems to directly influence learning and working memory capacity [26,27]. There may also be individual differences corresponding to working memory capacity in the maximum amount of associative activation from contextual cues, making base strength especially important for students with a smaller working memory and when complexity increases [26–29].

Students will ultimately be tested on their ability to solve problems, not their memory of physics principles and definitions. However, declarative memory is essential as problem solving requires a lot of retrieval of physics principles, definitions, and solution strategies, especially during the planning phase [30]. Performance tends to deteriorate when a problem-solving task becomes more complex and most of the errors and failures seem to be due to misretrievals [31]. Hence, poor memory strength may be disguised as poor problem-solving skills, with resulting calls for ever more problem solving. The cost of failure is also high when solving problems [32], in large part because of lost time spent floundering.

Do students really lack knowledge of physics principles and definitions? And is it not better to just learn them through regular study? The results from phase 1 of the study, reported here, show that many students have severe lacks in knowledge of the most basic physics definitions and principles. As a student in one of our tests remarked: “I do not walk around remembering equations six weeks before the exam,” unknowingly referring to Newton’s second law. Some students do learn physics definitions and principles by retrieving during individual problem solving, as some students reported in our surveys. However, most novices, and especially the weaker students, tend to search for specific equations in textbooks or cramming sheets while solving problems [33–35]. Retrieval practice has the potential to narrow the gap between stronger and weaker students by making principles and their conditions of application more accessible and recognizable to weaker students. Gjerde *et al.* [15] found that retrieval practice of physics principles and their conditions significantly increased the probability that students mentioned conditions of application of principles while solving physics problems (odds ratio = 5.76).

In this paper we present a novel tool descriptively called the hierarchical principle structure for mechanics (HPSM) (see Supplemental Material [36] for the current version). HPSM contains the most relevant principles and definitions for an introductory mechanics course at a large university in Norway and was designed by the first author. It is hierarchical in the sense that principles are placed in a meaningful order according to central concepts, and whether they are from translational, rotational, or fluid mechanics. We use the word “principle” to refer to all the equations that are not mere definitions. To some degree, the organization of HPSM also reflects the textbook

for the course [37]. Rawson and Dunlosky [38] remarked that memorization should probably be constrained to key concepts that provide the foundation for further learning. Novice students, particularly the weaker students, also lack cohesion in their domain knowledge [39]. HPSM can help make clear what the essential principles and definitions are and help integrate domain knowledge that might seem fragmented to a novice student. Most physics students are familiar with cramming sheets, as it is a normal practice to allow a set number of handwritten sheets for exams. A quick internet search reveals numerous examples of physics cramming sheets, some from commercial actors but most made by students. However, these cramming sheets usually lack meaningful organization and rather reflect students’ effort to include every equation and some diagrams. Others, such as those in physics textbooks, are usually in the form of tables of constants, concepts, and equations. The novelty in our study is in the hierarchical structuring of HPSM, reduction to the essential principles and definitions, and in integrating retrieval practice of HPSM into regular lectures (phase 2).

As already mentioned, research on retrieval practice in physics is scant. This study can be viewed as an early step towards finding a role for retrieval practice in physics education. We introduce retrieval practice of HPSM in two phases, where we explore five research questions. In the first phase, we performed a randomized controlled trial to find (i) whether 70 min of retrieval practice of HPSM improves basic factual knowledge compared to studying problems and (ii) whether students lack knowledge of basic facts after the concepts have been introduced through lectures and problem solving. In the second phase, we implemented longitudinal retrieval practice in physics lectures to explore (iii) whether participation correlates with basic factual knowledge or (iv) with exam results. We also collected exam results in phase 2 and could therefore answer (v) whether scores on basic factual knowledge correlate with exam scores.

II. METHODS

This study took place over two semesters in a calculus-based introductory mechanics course at the University of Bergen. The course participants came from a mixture of disciplines (Physics, Teacher Education, Nano Technology, Ocean Technology, Energy, and Petroleum Technology). Most have completed two years of physics at the high school level.

Phase 1 of the study was a randomized controlled trial comparing retrieval practice (intervention group) with studying physics problems (control group). We expected large effects of retrieval practice on a test of factual knowledge compared to control. A power analysis using GPower [40] with a large effect size of $d = 0.8$ and alpha level of 5%, suggested a minimum of 42 participants to achieve a power of 80% for detecting an effect.

The experiment was conducted during a regular lecture and participation was voluntary and anonymous. 81 students showed up out of the roughly 150 students signed up for the course. All participants had equal chances of winning a gift certificate of 2000 NOK (~250 usd) at the end of the experiment. All the concepts in the course curriculum had been covered in lectures before the experiment took place. The intervention group did 70 min of written retrieval practice of the hierarchical principle structure for mechanics (see Supplemental Material [36] for the current version of HPSM) on a worksheet where parts of the HPSM were removed. The worksheet consisted of 8 pages where parts of the HPSM had been progressively removed. The students received instructions to retrieve from memory and write down the missing parts. Some parts of HPSM were marked as not relevant (kinematics and fluid mechanics). The control group studied nine pairs of problems from the similarity judgment task (Appendix 1 in Ref. [41]). The nine problems covered the same concepts as the retrieval practice. The full HPSM was available during practice for both groups. Students in both groups had 20 min to first complete a filler post-test after completing the practice phase to get a better measure of long-term memory. Then, the students had 10 min to complete the declarative facts test which consisted of 20 questions.

What is/are the unit(s) of

1. Force
2. Energy
3. Work
4. Linear momentum
5. Angular momentum
6. Torque

Write an expression for

7. Newton's second law
8. Work when force is constant
9. Conservation of mechanical energy
10. Conservation of energy, non-conservative forces

included

11. The work-energy theorem.
12. The impulse-momentum theorem
13. Linear momentum
14. The angular momentum of a particle
15. The angular momentum of a rotating object
16. Gravitational potential energy
17. Spring potential energy
18. The force of friction

What are the conditions for

19. Conservation of linear momentum
20. Conservation of angular momentum

The first author constructed the test. The three authors, who all have at least five years of physics at university level and who all have taught physics, agreed that the test questions probe essential basic facts from mechanics. Cronbach's alpha for these 20 items was 0.89 in phase 1 and 0.87 in phase 2, indicating good internal consistency.

Phase 2 of the study was correlational, and was in part motivated by the results in phase 1 and promising results in Ref. [15]. We implemented distributed retrieval practice in the lectures of the same course as in phase 1, but in the subsequent year. The students completed 15 min of retrieval practice in the beginning of 16 of the lectures. Apart from the retrieval practice, the lectures mostly consisted of traditional lecturing and some weekly quizzes with conceptual questions and a peer instruction format [42]. Our participants were those students who decided to show up for lectures, which were not mandatory. In total, 130 students participated in retrieval practice at least once. The study sample in phase 2 consisted of approximately 35% females, 65% males, and 21% nano technology, 18% ocean technology, 13% physics, 13% energy, 9% teacher education, 7% petroleum technology, and 19% other, with a mean age of 21. The average show up was 53 students ($SD = 24$), roughly reflecting how many typically show up for voluntary lectures at the institute. Each written work (retrieval sheet) handed in counted as one lot for a lottery of three gift certificates (~110\$ each). The students could participate in all activities regardless of whether they chose to hand in their written material for analysis. The retrieval practice was performed on two-sided retrieval sheets where the equations, their names, and their conditions of application were removed from HPSM. The full HPSM was available during retrieval practice for feedback and restudy opportunity. Advice for how to do the retrieval practice was visible on a projector screen while the students practiced (see Supplemental Material [36] for the advice given to students and the literature the advice was based on). We tested the students on the same declarative facts test as in phase 1, but 37 days earlier in the semester. Furthermore, in phase 2 we obtained final exam scores for 90 students and prior Calculus 1 grades as a measure of prior ability for 83 of these 90 students. The final exam consisted of regular word problems and a few conceptual multiple-choice questions. We obtained exam results for 28 of the 34 students that completed the declarative facts test.

The data from the RCT in phase 1 was analyzed with a t test, with Cohen's d as a measure of the effect size. The data from phase 2 was analyzed with simple correlational analysis and regression. An important confounder in phase 2 is the fact that students also participated in the lectures when they did retrieval practice. Therefore, any correlations with performance in phase 2 needs to be interpreted with caution.

III. RESULTS

A. Phase 1—RCT

For reference, Cohen's d effect sizes of about 0.20, 0.50, and 0.80 are usually treated as small, medium, and large, respectively [43]. Hattie [44] proposes that effect sizes of 0.40 or higher are educationally relevant, although one must also consider ease of implementation.

A two-tailed t test was performed to determine whether retrieval practice significantly affected basic factual knowledge compared to problem study. The t test showed a significant effect of retrieval practice for score on the declarative facts test ($M = 14.6$, $SD = 3.6$) compared to control ($M = 9.0$, $SD = 4.3$), $t(77.5) = 6.4$, $p < 0.001$, $d = 1.42$, a very large effect.

B. Phase 2—Correlational

Roughly a third of psychological meta studies have $r < 0.20$, the middle third has r of 0.20–0.30, and the upper third have $r > 0.30$ [45,46]. We use these numbers as empirical guidelines for small, medium, and large effect sizes.

We calculated the Pearson correlation coefficients to find whether number of retrieval practice sessions attended correlated with score on the facts test and final exam score, and whether the facts test correlated with final exam score (research question 3, 4, and 5). All correlations were significant, see Table I (see Figs. 4, 5, and Fig. S6 in the Supplemental Material [36] for scatter plots). Moreover, a scatter plot of retrieval practice sessions vs exam score suggested a possible interaction between attending retrieval practice sessions and ability, where high-ability students do well regardless of whether they attend lectures with retrieval practice. A multiple linear regression was therefore calculated to predict the physics exam grade with Calculus 1 exam grade, the number of retrieval sessions attended, and the interaction as predictors. A significant regression equation was found [$F(3, 79) = 18.73$, $p < 0.001$], with an $R^2 = 0.42$ and adjusted $R^2 = 0.39$, see Table II. Both main effects were significant, but more importantly the interaction term was significant. In other words, weak students seem to benefit while strong students do well regardless.

IV. DISCUSSION

Answering research question 1, the RCT in phase 1 showed that 70 min of retrieval practice can be far more efficient than “just studying” for learning physics facts ($d = 1.42$). That retrieval practice is better than studying for a factual test is not surprising, but the effect is remarkably large. As implementation is easy, it seems ready for use in physics education. In answering research question 2—whether students actually lack knowledge of basic physics facts—we qualitatively evaluate the control group’s mean score of 9 correct answers to be very low when considering the low difficulty of the test items. Physics students lack the most basic knowledge.

In phase 2, the very high correlation ($r = 0.62$) between score on the factual test and the final exam suggests that knowledge of basic facts is important in introductory mechanics (research question 5). Participation in lectures with retrieval practice had a high correlation with score on the factual test ($r = 0.44$), which suggests that retrieval

TABLE I. Correlations between retrieval practice sessions attended, the declarative facts test, and final exam score.

Relationship	r	t value	df	p
Retrieval practice—facts test score	0.44	2.74	32	<0.01
Retrieval practice—exam score	0.33	3.29	88	<0.01
Facts test score—exam score	0.62	4.06	26	<0.001

practice also affects basic factual knowledge when distributed throughout the semester (research question 3). Still, something goes wrong for some of the students. As an example, one of the students attended seven retrieval practice sessions and still only got two factual questions correct (question 1 and 2) and there were others with similar results. It seems likely that these students either participated to win money, misunderstood the purpose and merely copied equations, or ignored advice.

The correlations in Table I and the regression model in Table II indicate that weaker students (as measured by Calculus 1 grade) benefit more than strong students from lectures with retrieval, while strong students stay strong (research question 4). Retrieval practice of principles and definitions may be an effective way to reduce the gap between weak and strong students by strengthening the weak.

We speculate that some elaboration, and possibly time for consolidation, is essential for getting maximum benefits from retrieval practice [21]. Many students retrieved principles without any prior elaborative encoding in phase 2, and some students did complain about lack of understanding of retrieved principles. We thought that students would try to encode the principles elaboratively without specific support, but it seems that some support is needed.

In hindsight, we probably made it unnecessarily difficult to memorize HPSM in phase 2 by removing every cue except spatial location. Vaughn and Rawson [47] found that cued retrieval practice enhances memory for both the memory cue and the target memory, enhancing associative memory in both directions although slightly more in the practiced direction. This might justify doing only forward recall of equations from name, conditions, and location. Retrieval of the equation from the name and the condition of application for the principle is also more aligned with what happens in problem-solving situations, which is

TABLE II. Multiple linear regression of Calculus 1 grade and retrieval sessions attended as predictors of exam grade, $N = 82$.

	ΔR^2	B	SE B	p
Step 1	0.42			
Constant		1.65	0.54	<0.01
Calculus 1 grade		0.71	0.14	<0.001
Retrieval sessions		0.17	0.08	<0.05
Calculus*retrieval		−0.04	0.02	<0.05

important for transfer to occur [25]. Students may also become more motivated to use retrieval practice as they experience quicker success [38]. We also constrained the students to retrieval of current concepts, with some opportunity for repetition of prior concepts. Providing the full HPSM on both sides of the retrieval sheet grants the student greater flexibility in what to study and retrieve, and probably improves learning of spatial locations (see Supplemental Material [36] for the current version of the retrieval sheets).

V. CONCLUSIONS

In phase 1 of this study, we showed that (control) students in introductory mechanics have a grave lack in basic factual knowledge, and that a short intervention with retrieval practice of a hierarchical principle structure can dramatically increase scores on a declarative facts test. In phase 2, we implemented distributed retrieval practice in

lectures throughout the semester. Results indicate that knowledge of basic facts predicts exam score, and that especially the weaker students benefit from attending lectures with retrieval practice.

More research is needed on how to better integrate the hierarchical principle structure into a course and how to support elaborative encoding, which will probably potentiate the effects of retrieval practice [21]. There is also a need for experimental testing of whether lectures with distributed retrieval practice is superior to lectures without, and for what measures.

We speculate that HPSM has a potential—beyond retrieval practice—as an organizing tool for lectures and other learning activities. However, testing this would require a different theoretical framework and research design.

ACKNOWLEDGMENTS

This work was supported by the University of Bergen.

-
- [1] J. Dunlosky, K. A. Rawson, E. J. Marsh, M. J. Nathan, and D. T. Willingham, Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology, *Psychol. Sci. Publ. Interest* **14**, 4 (2013).
 - [2] P. L. Pirolli and J. R. Anderson, The role of practice in fact retrieval, *J. Exper. Psychol. Learning Memory Cogn.* **11**, 136 (1985).
 - [3] R. L. Bangertdrowns, J. A. Kulik, and C. L. C. Kulik, Effects of frequent classroom testing, *J. Educ. Res.* **85**, 89 (1991).
 - [4] A. C. Butler, Repeated testing produces superior transfer of learning relative to repeated studying, *J. Exper. Psychol. Learning Memory Cogn.* **36**, 1118 (2010).
 - [5] J. D. Karpicke, Retrieval-based learning: Active retrieval promotes meaningful learning, *Curr. Dir. Psychol. Sci.* **21**, 157 (2012).
 - [6] H. S. Lee and D. Ahn, Testing prepares students to learn better: The forward effect of testing in category learning, *J. Educ. Psychol.* **110**, 203 (2018).
 - [7] B. Pastotter *et al.*, Retrieval during learning facilitates subsequent memory encoding, *J. Exper. Psychol. Learning Memory Cogn.* **37**, 287 (2011).
 - [8] H. L. Roediger and J. D. Karpicke, The power of testing memory: Basic research and implications for educational practice, *Perspectives Psychol. Sci.* **1**, 181 (2006).
 - [9] H. L. Roediger and J. D. Karpicke, Test-enhanced learning—Taking memory tests improves long-term retention, *Psychol. Sci.* **17**, 249 (2006).
 - [10] H. L. Roediger and M. A. Pyc, Inexpensive techniques to improve education: Applying cognitive psychology to enhance educational practice, *J. Appl. Res. Memory Cogn.* **1**, 242 (2012).
 - [11] T. Zu, J. Munsell, and N. S. Rebello, Comparing retrieval-based practice and peer instruction in physics learning, *Phys. Rev. Phys. Educ. Res.* **15**, 010105 (2019).
 - [12] C. A. Rowland, The effect of testing versus restudy on retention: A meta-analytic review of the testing effect, *Psychol. Bull.* **140**, 1432 (2014).
 - [13] J. D. Karpicke and P. J. Grimaldi, Retrieval-based learning: A perspective for enhancing meaningful learning, *Educ. Psychol. Rev.* **24**, 401 (2012).
 - [14] M. A. B. Whitaker, The role of memorization in physics education, *Am. J. Phys.* **53**, 111 (1985).
 - [15] V. Gjerde, B. Holst, and S. D. Kolstø, Problem solving in basic physics: Effective self-explanations based on four elements (to be published).
 - [16] J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, and Y. Qin, An integrated theory of the mind, *Psychol. Rev.* **111**, 1036 (2004).
 - [17] J. R. Anderson and C. Lebiere, *The Atomic Components of Thought* (Lawrence Erlbaum Associates, Mahwah, N.J. 1998), p. 490.
 - [18] J. R. Anderson, J. M. Fincham, and S. Douglass, Practice and retention: A unifying analysis, *J. Exper. Psychol. Learning Memory Cogn.* **25**, 1120 (1999).
 - [19] N. Kornell, R. A. Bjork, and M. A. Garcia, Why tests appear to prevent forgetting: A distribution-based bifurcation model, *J. Memory Lang.* **65**, 85 (2011).
 - [20] J. R. Anderson, *The Adaptive Character of Thought. Studies in Cognition* (L. Erlbaum Associates, Hillsdale, NJ, 1990), p. 276.
 - [21] J. W. Antony, C. S. Ferreira, K. A. Norman, and M. Wimber, Retrieval as a fast route to memory consolidation, *Trends Cognit. Sci.* **21**, 573 (2017).

- [22] S. K. Carpenter, Testing enhances the transfer of learning, *Curr. Dir. Psychol. Sci.* **21**, 279 (2012).
- [23] B. Pastotter and K. H. Bauml, Retrieval practice enhances new learning: The forward effect of testing, *Front. Psychol.* **5**, 1 (2014).
- [24] D. Rohrer, K. Taylor, and B. Sholar, Tests enhance the transfer of learning, *J. Exper. Psychol. Learning Memory Cogn.* **36**, 233 (2010).
- [25] S. C. Pan and T. C. Rickard, Transfer of test-enhanced learning: Meta-analytic review and synthesis, *Psychol. Bull.* **144**, 710 (2018).
- [26] L. M. Reder, X. L. Liu, A. Keinath, and V. Popov, Building knowledge requires bricks, not sand: The critical role of familiar constituents in learning, *Psychon. Bull. Rev.* **23**, 271 (2016).
- [27] Z. F. Shen *et al.*, Item strength affects working memory capacity, *Memory Cogn.* **46**, 204 (2018).
- [28] P. K. Agarwal *et al.*, Benefits from retrieval practice are greater for students with lower working memory capacity, *Memory* **25**, 764 (2017).
- [29] M. C. Lovett, L. Z. Daily, and L. M. Reder, A source activation theory of working memory: cross-task prediction of performance in ACT-R, *Cogn. Syst. Res.* **1**, 99 (2000).
- [30] J. R. Anderson and J. M. Fincham, Extending problem-solving procedures through reflection, *Cogn. Psychol.* **74**, 1 (2014).
- [31] J. R. Anderson, L. M. Reder, and C. Lebiere, Working memory: Activation limitations on retrieval, *Cogn. Psychol.* **30**, 221 (1996).
- [32] P. I. Pavlik and J. R. Anderson, Using a model to compute the optimal schedule of practice, *J. Exp. Psychol. Appl.* **14**, 101 (2008).
- [33] J. Larkin, *The Role of Problem Representation in Physics, Mental Models Conference* (University of California, San Diego 1981).
- [34] W. J. Leonard, R. J. Dufresne, and J. P. Mestre, Using qualitative problem-solving strategies to highlight the role of conceptual knowledge in solving problems, *Am. J. Phys.* **64**, 1495 (1996).
- [35] A. J. Mason and C. Singh, Surveying college introductory physics students' attitudes and approaches to problem solving, *Eur. J. Phys.* **37**, 055704 (2016).
- [36] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevPhysEducRes.16.013103> contains the hierarchical principle structure, a retrieval practice worksheet, scatter plots, advice for retrieval practice, and questions for elaborative encoding of principles.
- [37] H. D. Young, R. A. Freedman, and A. L. Ford, *Sears and Zemansky's University Physics: With Modern Physics* (Pearson Education, Essex [England], 2016).
- [38] K. A. Rawson and J. Dunlosky, When is practice testing most effective for improving the durability and efficiency of student learning?, *Educ. Psychol. Rev.* **24**, 419 (2012).
- [39] P. A. Alexander, The development of expertise: The journey from acclimation to proficiency, *Educ. Res.* **32**, 10 (2003).
- [40] E. Erdfelder, F. Faul, and A. Buchner, GPOWER: A general power analysis program. *Behavior Research Methods, Instrum. Computers* **28**, 1 (1996).
- [41] P. T. Hardiman, R. Dufresne, and J. P. Mestre, The relation between problem categorization and problem-solving among experts and novices, *Memory Cogn.* **17**, 627 (1989).
- [42] E. Mazur, *Peer Instruction: A User's Manual. Prentice Hall Series in Educational Innovation* (Prentice Hall, Upper Saddle River, NJ, 1997), p. 253.
- [43] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. (L. Erlbaum Associates, Hillsdale, NJ, 1988), p. 567.
- [44] J. Hattie, *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement* (Routledge, London, New York, 2009), p. 378.
- [45] J. F. Hemphill, Interpreting the magnitudes of correlation coefficients, *Am. Psychol.* **58**, 78 (2003).
- [46] G. E. Gignac and E. T. Szodorai, Effect size guidelines for individual differences researchers, *Personality Individ. Diff.* **102**, 74 (2016).
- [47] K. E. Vaughn and K. A. Rawson, Diagnosing criterion-level effects on memory: What aspects of memory are enhanced by repeated retrieval?, *Psychol. Sci.* **22**, 1127 (2011).