Increasing the effectiveness of active learning using deliberate practice: A homework transformation

Kelly Miller,[1] Kristina Callaghan,[2] Logan S. McCarty,[2,3] and Louis Deslauriers[2,*]

[1]*School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA*
[2]*Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA*
[3]*Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, USA*

We show how learning can be improved, beyond that shown in actively taught classrooms, by also transforming the homework using the principles of deliberate practice. We measure the impact of transforming the homework on student learning in a course that had already implemented an active approach to teaching in class. We compare performance on the same final exam in equivalent cohorts of students over three semesters of an introductory physics course: the first taught with traditional lectures and traditional homework, the second taught with active instruction coupled with traditional homework, and the last taught with both active instruction and transformed homework. We find students in the semester where both active teaching and transformed homework are used scored significantly higher on the final exam than the students taught actively but with traditional homework. This learning gain achieved by transforming the homework is comparable to that achieved by replacing traditional lectures with active teaching strategies in class. We further show the positive effects of transforming the homework on student learning through a shorter, controlled experiment. When everything but the homework implementation is controlled, students scored 5%–10% higher on a test of learning following transformed homework compared with traditional homework. This significant improvement to learning occurs despite students spending a similar amount of time on task. This study represents an initial step towards understanding how deliberate practice can be extended to improve pedagogy beyond what happens in the classroom to the out-of-class homework.

DOI: 10.1103/PhysRevPhysEducRes.17.010129

I. INTRODUCTION

Active learning pedagogies are the "gold standard" of teaching. Extensive research shows that students learn more when they are actively engaged in the classroom than they do in a passive lecture environment [1–6]. One of the pedagogical components of active teaching that has been shown to foster successful learning is "deliberate practice" [7–10]. As described by Ericsson [7], these principles are fundamental in determining how people learn complex tasks. By reviewing research on skill acquisition in many disparate domains including chess, musical performance, and athletics, Ericsson identified specific elements that are key to developing expertise in a field. These elements, collectively known as deliberate practice, include the following: First, an expert tutor or coach uses their expertise to break down a complex task or skill into well-defined "subskills." The learner then practices these subskills repeatedly while the coach provides targeted, expert feedback. Through repetition and feedback, the learner gradually refines and improves their performance to achieve appropriate mastery of the subskills. Finally, the expert coach designs exercises for the learner to synthesize the subskills in order to practice, and receive feedback on, the complex target task. The tutor or coach plays three essential roles: identifying and designing the subskill practice, giving expert feedback and motivation, and helping to integrate the subskills to achieve the complex goal. While feedback from a coach and an emphasis on subskills are essential for deliberate practice, equally important is the role of the learner's engagement and willingness to devote focused, effortful practice. As a learner gains expertise, they can increase their self-efficacy and become better at evaluating their own performance [7,8].

Previous studies have shown successful implementation of deliberate practice to transform teaching strategies in the classroom and improve student learning [3,6,11,12]. Typically, these studies have changed the in-class interaction between the teacher and the students but have not investigated the possibility of additional learning gains

---

*To whom correspondence should be addressed.
louis@physics.harvard.edu

from applying deliberate practice to the work done outside of class. Homework is implicitly a form of active learning, but traditional homework does not incorporate the principles of deliberate practice such as subskill practice with targeted feedback. Instead, most homework involves complex problems similar to the questions that will be asked on exams.

While most homework pays little or no attention to deliberate practice, some online computer homework or tutoring systems have emerged with varying levels of subskill development and instant feedback. Many of these systems can be useful instructional tools and advantageous for student learning and improvement [13–17]. Mikula and Heckler [13] found that students' lack of basic skills (like vector addition) presents a serious obstacle to learning physics and solving more complex problems [17]. Using computer-based instruction, immediate feedback, mastery grading, interleaved training sequences, and distributed practice, they developed an instructional framework to address specific student difficulties with these basic skills.

Many of these online systems integrate the principles of "mastery-style homework." "Learning for mastery" was an idea first expressed by Bloom in 1968 [18] who proposed that learning should be paced and structured so that students at different levels of mastery are exposed to different instructional pathways or treatments. Students who are struggling with specific concepts, for example, receive feedback and treatment as well as spend more time until they master that content. There are similarities between mastery-style homework and homework designed with the principles of deliberative practice. For instance, both focus on breaking down learning goals into subskills and ensuring complete mastery before moving on to the next skill. However, deliberate practice requires optimizing the specific subskills for an individual student or particular population, and emphasizes the integration of subskills towards the goal of performing more complex tasks all the while providing students with targeted and timely feedback as they progress toward mastery.

Thus, in deliberate practice, an expert tutor is needed to identify the *optimal set* of subskills that can provide the most benefit to a particular student population and write hints for each subskill that are consistent with student thinking for that population. This requires expert knowledge of student thinking for that population, and consequently, most online homework systems and computer tutors cannot do this. Second, the subskill exercises we focus on in this homework transformation are distinct from traditional online homework systems because they are immediately followed by more complex problems that integrate the subskills.

In this paper we address the question of whether applying the principles of deliberate practice to homework can increase learning beyond what is routinely achieved with the existing "active learning" gold standard. We compare student learning in actively taught physics courses before and after the homework transformations. We show that the learning gains achieved with these homework transformations are comparable to the learning gains achieved when transforming lectures from a traditional (passive) to an active instructional format.

II. METHODS

This investigation was conducted in three different large-enrollment introductory physics courses at Harvard University and featured two separate *course-wide transformations* and one *targeted experiment*. The two course-wide transformations served as quasiexperimental studies of the impact of a homework transformation over an entire semester, while the targeted experiment measured learning gains using a controlled randomized design.

Before the start of this investigation, all three courses had traditional-style homework assignments. Each semester-long course had nine assignments, each covering one "course module" (2–4 lectures). Each assignment had an online portion of 10 multistep exam-style problems plus an offline portion of about five similarly complex problems that students wrote out and submitted on paper. The paper questions were graded by teaching assistants and returned a week later, while the online questions were graded correct or incorrect by *Sapling* [19], the online homework system we used. Sapling is the online homework platform offered through Macmillan Learning. It allows instructors to create homework assignments from a database of questions or from their own questions which they enter into the system. Students completed and submitted the homework online and were able to revise and resubmit their answers as many times as they wanted but lost 5% of the total question credit for each incorrect answer submitted. This small deduction ensured a reasonable effort on the part of the students and discouraged students from simply guessing at answers. Other than learning that their answers were correct or incorrect, students did not receive additional feedback or hints with these online questions.

Each transformed homework consisted of the same set of online exam-style problems and offline (paper) problems *plus* an additional 25–30 online subskill questions. Students were instructed to complete all of the subskill problems before attempting any of the more complex problems. To design each transformed homework, an expert tutor identified approximately 10 specific learning goals and wrote 2–4 subskill questions targeting each learning goal. As we will discuss below, the question author must be well versed in common student difficulties and misconceptions—content knowledge alone is woefully insufficient. Each subskill question had two or more targeted hints that students could view if they were stuck, and a detailed solution would appear once a student submitted a correct answer or "gave up" on a problem. (Students could resubmit incorrect answers with the same

TABLE I.   Comparison of traditional versus transformed homework implementation.

	Traditional homework	Transformed homework
Content	10 online, "examlike" problems *plus* 5 similarly complex paper problems	25–30 subskill questions *plus* same 10 online, "examlike" problems *plus* 5 similarly complex paper problems
Feedback	Online problems: correct or incorrect paper problems: feedback delayed by a week	Immediate feedback on subskills *plus* progression of targeted hints *plus* detailed solution when right answer is submitted or student "gives up" *plus* usual feedback on traditional q's.

5% penalty as the other online problems.) The hints and explanations were designed to give progressively more assistance to a confused student. Since each question targeted a narrow subskill, the expert tutor could write hints that addressed all of the most important misconceptions. In contrast, it is extremely difficult to write effective online hints for complex multistep problems because these problems have so many different ways that students can get them incorrect. Examples of subskill problems with a progressive series of hints are shown in the Appendix. Table I summarizes the differences between the traditional and transformed homework.

A. Context of the course-wide transformations

The two course-wide transformations took place in separate calculus-based introductory physics courses at Harvard University. These two courses, Physical Sciences 2 (PS2) and Physical Sciences 3 (PS3), comprise a one-year sequence aimed at life sciences and pre-medical students, with topics ranging from Newtonian mechanics, fluids, and statistical physics in the fall term (PS2), to electricity, magnetism, waves, and optics in the spring (PS3). With a typical enrollment of 180 to 220 students, these are the largest courses offered each year by the physics department. Before the study took place, the courses were already well established with an extensive set of lecture notes, tutorial material, weekly homework, review materials, and laboratories. These courses earned consistently strong student evaluations and were led by experienced faculty who taught in a traditional lecture style. PS2 and PS3 attract the majority of the pre-medical and life science students who take physics at Harvard, so the enrollees are quite similar from year to year—they have similar scores on college entrance exams, similar prior coursework, and similar motivation for taking these courses. In fact, in previous years we had routinely administered the Force Concept Inventory (FCI) [20] as pre- and post-tests, but the data were so similar from year to year that we had stopped that practice before the years described in this study. With this remarkably stable student population, these courses are ideal for quasiexperimental studies across cohorts.

The course transformation in PS2 took place in two stages across three cohorts of students. In the first year of the study, the students in this initial cohort (C1, $N_{PS2} = 208$) received traditional lectures and homework. This group served as the baseline control for the interventions that took place in the second and third year of the study. The traditional lectures for C1 consisted of a mix of chalkboard and PowerPoint presentations enhanced with elaborate physics demonstrations [21], and the traditional homework assignments were structured as described above. In the second year of the study, the lectures for the students in cohort C2 ($N_{PS2} = 189$) were transformed with the use of deliberate practice principles to reflect the active learning strategies used in Refs. [3,6,11,12,22]. The homework for C2 retained its traditional format. Finally, in the third year, the students in cohort C3 ($N_{PS2} = 198$) experienced transformed lecture and homework simultaneously. With this implementation we could examine the effect of adding homework based on deliberate practice to a course that had already been transformed to use active learning in the classroom.

The course transformation in PS3, by contrast, took place all at once. In the first year, students in cohort C1 ($N_{PS3} = 200$) had the baseline experience with traditional lectures and homework. Then, students in cohort C2 ($N_{PS3} = 190$) received the full treatment with active learning in the lecture and deliberate practice on the homework. Table II summarizes the type of pedagogy used in the lecture and homework components in each year for both courses as well as which of three instructors (A, B, C or some combination) taught each course. Note that the students in cohort C2 experienced active learning with traditional homework in PS2, followed immediately by active learning with deliberate practice homework in PS3. Thus, this group of students experienced both types of homework within a consistent pedagogical framework of active learning. Given that PS2 was offered in the fall semester and PS3 in the spring semester, there was some overlap of students between the cohorts in this study. Out of the 208 students enrolled in PS2 in year 1 (C1), 172 (~85%) were also in PS3 in year 1. Thirty-six students in PS2 C1 did not move on to PS3 C1 ("drop outs") and 28 students in PS3 C1 came in ("drop ins"). In year 2, 168 of the 189 students (~88%) enrolled in PS2 (C2) were also enrolled in PS3 in year 2. Twenty-one students in PS2 C2 did not move on to PS3 C2 ("drop outs") and 24 students in PS3 C2 came in ("drop ins").

TABLE II.   Pedagogical style in the lecture and homework portions for both courses, across three cohorts in three consecutive years.

Course	Year: Cohort	Instructor(s)	Enrollment	Lecture	Homework
PS2	Year 1: C1	A & C	208	Traditional	Traditional
	Year 2: C2	A & B	189	Active learning	Traditional
	Year 3: C3	A & B	198	Active learning	Transformed (deliberate practice)
PS3	Year 1: C1	A	200	Traditional	Traditional
	Year 2: C2	A	190	Active learning	Transformed (deliberate practice)

To measure conceptual understanding and content mastery in PS2, the same three-hour comprehensive final exam was administered at the end of the course in all three cohorts. This exam was never released to students during the three-year study, so it offered a fair comparison of student learning. We have evidence from other Harvard physics classes (including in PS2 in years prior to this study) that there is no improvement in identical exams administered repeatedly in the same class from one year to the next. We can therefore be reasonably confident that any variation in exam scores from one year to the next is the result of student performance and not due to compromised exams.

The exam had six multiple-choice conceptual questions of a style similar to those on the Force Concept Inventory [20], along with seven complex, open-ended problems that require higher order thinking. In PS3, the 3-h-long final exams given to the two cohorts were also comprehensive and consisted of eight conceptual multiple-choice questions similar to those found in the Brief Electricity and Magnetism Assessment [23], followed by eight complex, open-ended problems. Unlike the PS2 exams, which were identical from year to year, the exam overlap between year 1 and 2 of PS3 was not 100%. Out of 8 long answer problems, 5 were common in both years. Out of 8 short multiple-choice questions, 4 were common in both years. In terms of point total, ~60% of the problems on the year 1 final exam were repeated in year 2. Most importantly, the amount of coverage for each topic was the same across both years. For instance, the *RC* discharging circuit problem that was dropped in year 1 was replaced by another *RC* discharge topic in year 2. To reassure ourselves that students had not gotten hold of the earlier exam, we compared students' performance on the identical problems to their performance on those that were different. There was no meaningful difference in performance between these two groups of questions. For the purpose of this study, we included only the scores on the identical questions from the PS3 exams. At the conclusion of the three-year study, all exams were regraded by a small team of graduate students, using the same metric that was used each semester in the previous years. These graders were blind to the cohorts of each exam.

As the transformed homework was considerably longer than the traditional homework, with 25–30 subskill questions added on top of all of the complex problems, we sought to understand the time required for students to complete each type of homework. We hypothesized that increased learning from the subskill questions could help students complete the complex problems more quickly, so the total time-on-task might increase only modestly despite the large increase in length of the transformed assignments. Thus, in addition to measuring student learning with final exam performance, an end-of-semester survey was used to measure the average time students spent each week completing the homework. Students were asked to state the average number of hours per week spent on the course outside of class. Additional survey questions revealed that students in these courses typically spend the vast majority of their out-of-class time completing the weekly homework, so out-of-class time is a useful proxy for homework time.

B. Context of the targeted experiment

The two course-wide implementations provide valuable insight into the use of deliberate practice throughout an entire course, but these quasiexperimental interventions have limited value in isolating and quantifying the learning gains due to the transformed homework alone. In addition, students were not asked directly to report the time-on-task for the homework. Thus, a targeted experiment with randomized student groups and specific questions about homework time was implemented in a third introductory physics course, Physics E1ax, which is the largest enrollment physics course taught at the Harvard Extension School. The Harvard Extension School offers liberal arts and professional courses, academic certificates, undergraduate and graduate degrees, as well as a pre-medical program. The instruction in Physics E1ax was in-person (as opposed to online). This course is the first semester of a year-long algebra-based introductory physics sequence (E1ax and E1bx) designed to match the PS2 and PS3 courses in every aspect except for the removal of calculus: they share the same course content, same instructor, nearly identical lectures, similar homework and exams, and are even taught in the same lecture hall. The lectures in these courses were transformed from traditional lectures to active learning at the same time as PS2 and PS3, by the same instructor. The students in these extension school courses

are mostly post-baccalaureate students who are taking physics as a pre-medical or other pre-health requirement, plus a wide range of other students including advanced high school students, extension school undergraduates, and lifelong learners.

The experiment took place during two weeks (modules) of the course and involved $N = 153$ students. One module covered the topic of static equilibrium, while the other covered fluids. During these two weeks, the class was randomly divided into two groups. For the statics module, Group 1 completed the transformed version of the homework while Group 2 did the traditional homework. The transformed and traditional homework in this targeted experiment followed the same implementation as for the course-wide study (see Table I). For the fluids module, the treatments were switched: Group 1 received the traditional homework while Group 2 had the transformed homework. At the end of each module, after completing the homework, students completed a ten-item, multiple-choice test of learning (TOL) on the content covered in that module. Students were encouraged to try their best on each TOL and were told that they would be good practice for the final exam but were reminded that their score on the TOL would not directly affect their course grade. Students were also told that they would receive participation points toward their final grade for completing the TOL. The homework was due at the beginning of the lecture each week, and students were given the TOL during lecture on that same day. Some students might have finished their HW several days prior to the TOL while others might have finished it a few hours prior. The TOL questions are provided in the Supplemental Material [24]. In addition to measuring student learning, with the TOLs, we also kept track of the amount of time students spent on the homework. We measured time on task by asking students, as soon as they completed their homework, to state how many hours they had spent on the corresponding homework. This allowed a comparison of the time on task for the two homework treatments.
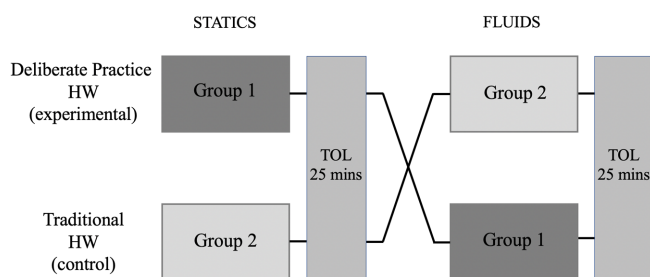

FIG. 1. Schematic showing the crossover study design of the targeted experiment. Students are randomly assigned to two groups and each student experiences both homework (HW) conditions across two physics topics (statics and fluids). A 25 min-long test of learning (TOL) is administered after each HW, for each topic.

Figure 1 summarizes the design for this targeted experiment. The study design featured a number of controls to ensure consistency and avoid bias: (i) Students were randomly assigned to the two groups. (ii) Each student experienced both types of homework in a "crossover" design that controls for other possible variation between students. (iii) The instructor did not see the TOLs, which were prepared independently by another author. (iv) The author of the TOLs wrote the tests based only on a list of detailed learning objectives for each module and did not have access to the course materials such as lectures or homework. (v) The authors of both versions of the homework (transformed and traditional) did not have access to the TOLs.

III. MEASUREMENTS AND RESULTS

This section shows the results of the two course-wide implementations, the targeted experiment and a follow-up experiment designed to further investigate the comparison of time on task between the two homework implementations.

A. Learning outcomes in course-wide transformations

Here we report on learning in the two course-wide transformations: the three-year progressive transformation in PS2 and the two-year transformation in PS3.

1. Course-wide transformation in PS2

The impact of transforming the homework on student learning is shown in Fig. 2(a) which compares student performance on the same 3-h-long comprehensive final exam over the three years of the study. Exam averages for C1, C2, and C3 were 70%, 77%, and 82%, respectively. Students in cohort C3, who experienced both active learning in lecture and the transformed homework, scored 5% better on the final exam compared with students in C2, who were taught actively but with traditional homework. Students in C2 scored 7% better on the final exam than students in C1, who were taught both with passive lectures and traditional homework. Thus, the learning gain achieved by transforming the homework was comparable to that achieved by replacing traditional lectures with active teaching strategies in class.

Figure 2(b) shows the results of students' responses to the question about out-of-class workload on the end of semester course evaluation. The average workload outside of class (in hours) is plotted for each of the three years of the study. Students in C1 and C2 (traditional homework) reported spending, on average, 6.9 h/week and 7.3 h/week, respectively, while students in C3 (transformed homework) reported spending 6.8 h/week. Although total time outside of class includes some activities unrelated to homework, and retrospective self-reports of time are notoriously inaccurate, it is notable that
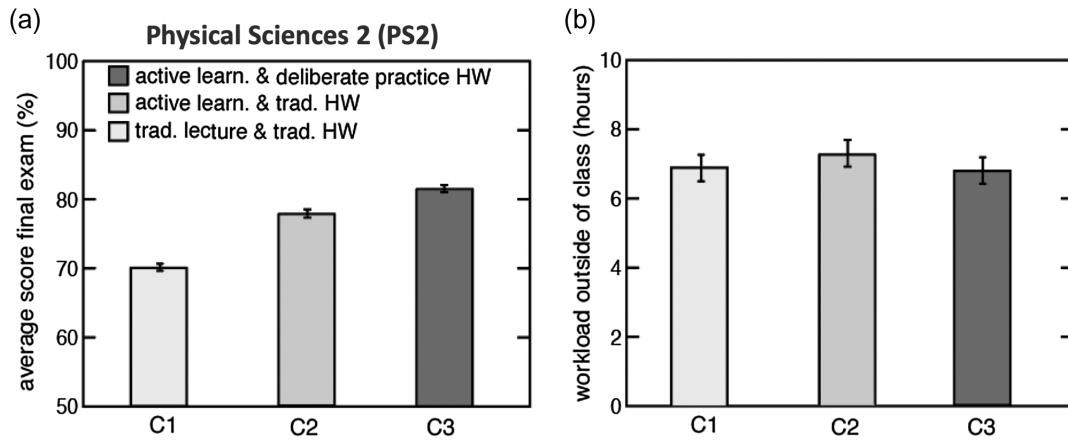
FIG. 2. Measurement of the (a) average percentage score for the 3-h comprehensive examination and (b) average workload (measured in hours) outside of class, for three cohorts of students over three consecutive years. The first cohort (C1) received a traditional treatment for the lecture and homework portions of the course. The second cohort (C2) experienced a change in lecture with the introduction of active learning but maintained the same traditional homework as C1. The third cohort (C3) received the transformed homework (deliberate practice HW) as well as the active learning in lectures. Error bars on both plots show 1 SE.

adding 25–30 extra subskill questions every week in C3 did not lead to a significant change in self-reported time spent out of class.

2. Course-wide transformation in PS3

In contrast to the progressive course transformation in PS2, the lecture and homework components of PS3 were transformed simultaneously between C1 and C2. As discussed in Sec II, we measured learning using a set of identical questions on the final examinations from these two years. These identical questions made up two-thirds of each exam, corresponding to roughly 2 h out of a 3-h exam, and offered comprehensive coverage of the topics in the course. As shown in Fig. 3(a), the scores on the identical

portions of the final exam went from 77% to 86% [see Fig. 3(a)]. This learning gain is nearly identical to the one found in the three-year study in PS2 (i.e., from C1 to C3). Figure 3(b) shows the results of students' responses to the question about out-of-class workload on the end of semester course evaluation. Figure 3(b) shows that the average workload increased significantly (but only marginally) between the two implementations of the homework. The average workload outside of class in C1 (traditional lecture and homework) was 5.8 h/week and 6.7 h in C2 (transformed lecture and homework). Although both lecture and homework were transformed, it is reasonable to conclude that this modest (but statistically significant) increase is due to the homework transformation, which added 25–30 subskill questions to each homework assignment.
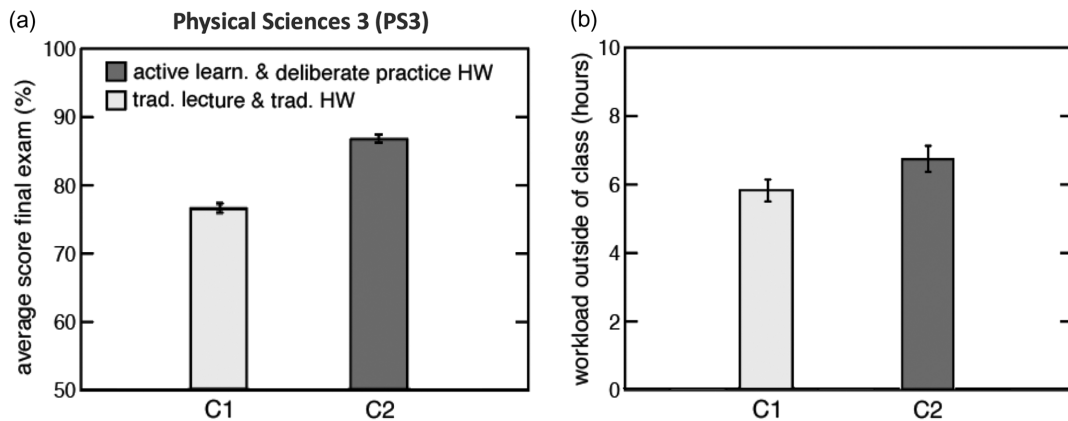

FIG. 3. (a) Measurement of the average percentage score for identical questions on the final examinations in PS3 for two cohorts of students over two consecutive years. The first cohort (C1) received a traditional treatment for the lecture and homework portions of the course. The second cohort (C2) experienced an entire course transformation, consisting of active learning in lecture and transformed homework (deliberate practice HW). (b) Student reported workload (outside of class, measured in hours). Error bars for both plots show 1 SE.
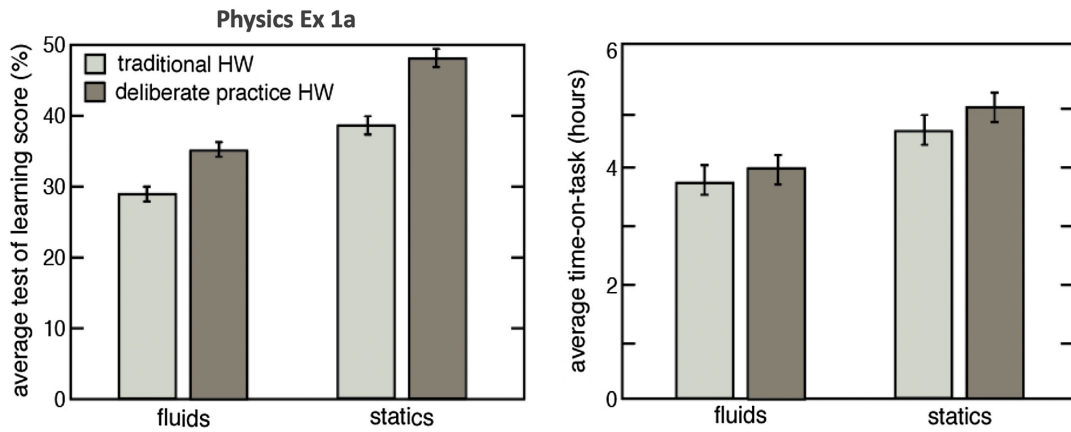
FIG. 4.   (a) Average performance on the tests of learning and (b) average time on task for students who completed the traditional homework (light gray) compared to those who completed the deliberate practice-based homework (dark gray). The same protocol was repeated for two physics topics (statics and fluids). Error bars on both plots show 1 SE.

The full-year experience of cohort C2 provides a unique perspective on the added benefits of transformed homework in a course with active learning. Following their fall course (PS2) with active learning in lecture and traditional homework, they continued in the spring semester (PS3) with active learning in lecture, but transformed homework. These students were able to compare directly the traditional and transformed homework in two courses taught by the same instructors. Student responses on an end-of-semester survey in PS3 were overwhelmingly positive. To the question "The homework in PS3 was designed differently from that in PS2. Did you find this new design helpful for your learning?" Students provided comments such as *"Yes. It really helped to get a foundation in the skills necessary to do the hardest problems," "Yes! Extremely. The ability to make sure I understood the different pieces before solving complex problems allowed me to know what parts I was missing." "Yes. I preferred doing more problems if it means that they first asked the basic questions that we were supposed to understand. I felt like in PS2 I spent hours trying to figure out very difficult problems that had many steps in them, and in that method the main point that was supposed to get across got lost in the process..."* On another survey during that same semester, PS3 students were asked a similar question: "The new structure of the homework in PS3 was helpful to my learning," to which they agreed strongly: 4.6 out 5 on a Likert scale.

B. Learning outcome of targeted experiment

For the targeted experiment in Physics E1ax, Figs. 4(a) and 4(b) highlight the difference in student learning and time on task between students in the traditional and transformed homework groups. We note, in particular, the following observations (all of which are confirmed by a more detailed statistical analysis): (i) Scores on the tests of learning are significantly higher in the transformed homework group. (ii) Time on task (the amount of time students report spending on homework) is similar between traditional and transformed homework. (iii) These trends are similar for the statics and fluids topics. Given the crossover study design, it appears that the improvement in TOL scores between traditional and transformed homework was not strongly affected by the choice of topic. It is important to point out that the traditional homework (and complex problems on the transformed homework) as well as the tests of learning were created from the learning objectives, by different authors. Therefore, the tests of learning were designed to be equally good measures of both the traditional homework and transformed homework.

We constructed a linear regression model (fixed effects) to identify the factors contributing to these observed differences in TOL scores. The model predicts students' performance on the TOL controlling for students' reported time on task and the homework content. Table III summarizes this model.

TABLE III.   Standardized coefficients for linear regression model predicting students' performance on the test of learning. Model controls for class topic (fluids versus statics), self-reported time spent on task, and two measures of background proficiency (scores on midterms and on the Colorado Learning Attitudes about Science Survey, CLASS [25]).

	Test of learning
Constant	-0.53***
Type of homework (traditional versus transformed)	0.46***
Topic (fluids versus statics)	0.68***
Average of first 2 midterms	0.48***
CLASS score	0.02
Time spent on task	-0.0004
R^2	0.17
RMSE	0.92

***$p < 0.001$.

The students in the transformed homework groups perform almost half a standard deviation better on the test of learning than the students in the traditional homework groups ($p < 0.001$). The crossover study design allowed us to control for any additional person-level variation by adding a categorical variable for each individual student; we found no meaningful change using these additional covariates. Conversely, removing all other student-level covariates (midterm average, CLASS score, and time on task) yielded identical results.

C. Miniexperiment on time on task

It is notable that, similar to the course-wide transformation in PS2, the self-reported time on task was not significantly greater despite adding 25–30 subskill questions to each assignment. Since self-reported assessments of time can be inaccurate, we conducted a follow-up experiment with a group of students in a subsequent semester in PS3. In this miniexperiment, students were asked to complete an abbreviated version of a problem set on the topic of dc circuits while supervised by a teaching assistant during a regular discussion section. There were 14 discussion sections, which were randomized at the cluster level: seven randomly chosen sections served as controls, while the others received the experimental treatment. Within a fixed 40-min time period, the control group completed two complex exam-style questions, while the experimental group spent 15 min on eight subskill questions followed by 25 min on the same two exam-style problems. The instructional staff controlled and monitored the time on task for both the control and experimental sections which appeared to be the same in both groups. At the end of the section, students were given a test of learning (scored out of 10 points) based on the concepts in the homework. Students in the transformed homework sections scored significantly higher on this test of learning compared with students in the traditional homework sections (93.7% compared to 80.7%, with $p < 0.001$). This suggests that students in the transformed homework groups learn more than the students in the traditional homework groups, despite spending similar time on task.

IV. DISCUSSION AND CONCLUSIONS

Through both a course-wide study and a targeted, controlled experiment, we have shown that students learn significantly more when the homework is transformed using the principles of deliberate practice, at no appreciable cost to the time spent on task. In the semester-long homework transformations, it is particularly noteworthy to see an effect on final exam performance given all the other factors that might influence student learning over an entire semester. The two-week implementation shows the more targeted impact of changing only the homework in a shorter, more tightly controlled experiment. In all three implementations, the increase in exam performance achieved through the transformed homework were 5%–10%. This is comparable to those achieved by transforming the lecture in year 2 of the study and consistent with the what is expected from this type of semester-long lecture transformation [4,11,12,22]. This indicates that instructors could gain as much by transforming the out-of-class component as they do by implementing active teaching in class.

To our knowledge, this is the first study to measure learning gains from applying the principles of deliberate practice to the homework for a class that had already optimized the in-class instruction through the implementation of these principles with active learning. Literature has shown that students learn more when homework is implemented through an online system instead of on paper [26–28]. It is worth noting that even before the principles of deliberate practice were used to transform the homework, the course was already using an online homework system which allowed students to revise and resubmit answers. Since the traditional homework in this study had already been optimized with an interactive online component, the change introduced by deliberate practice must be due to the design of those homework questions rather than the online format itself. A survey of the existing literature on deliberate practice suggests at least three factors that could contribute to the improved learning gains for the students in the transformed homework groups: (i) explicit identification of key knowledge-type subskills [29], (ii) repetition with targeted feedback, and (iii) assimilation via practice of the complex task [7–10,30,31]. The fact that students learn significantly more after completing the transformed homework illustrates the power of practicing subskills, having the opportunity to practice assimilation of subskills through the complex problems, and receiving targeted feedback throughout the process.

We have successfully implemented this type of homework transformation in a dozen courses in physics, chemistry and engineering at Harvard and four physics courses at the University of California, Merced. Student evaluations for these courses have shown that students enjoy the transformed homework implementation significantly more than traditional homework. Students indicate in these evaluations that they viewed the transformed homework as an essential tool for succeeding in the course. For example, in an introductory physics course for engineers at Harvard, students who experienced deliberate practice homework were asked to rate their level of agreement to the following questions on an end of semester survey: (i) "The subskill portion of each homework should remain part of the course next year", and (ii) "I wish the weekly homework in my other science courses (e.g., chemistry, math) would also develop subskills as part of their homework." Students answered 4.6 and 4.1 on a 5-point Likert scale for survey questions (i) and (ii), respectively. From the instructor's point of view, the transformed homework is easy to administer and grade, and leads to increased learning gains. These features have led instructors to view the implementation of

transformed homework as a "no brainer" and indicated that this way of doing homework would remain a consistent part of these courses in future offerings.

Given that this transformed homework implementation improves student learning with no appreciable difference in time on task and given that both instructors and students love it, one might wonder why all courses at Harvard have not adopted this style of homework. Successful implementation relies on having expert tutors to develop the subskill questions [32,33], but not every course has access to such personnel. These tutors must not only have experience teaching the course, but also must have spent a considerable amount of time talking to students about the concepts and developed an in-depth understanding of student thinking and common misconceptions. We have seen occasional examples of homework transformations that have been rejected by students. In one case, involving a physics course at Harvard, we had a content expert writing the subskill questions, but that person did not have expert knowledge of students' thinking. The resulting subskill questions took too long to answer and did not target the right concepts to help students synthesize their knowledge when doing the complex problems. Effective subskill questions must be written in such a way that they can be answered very quickly and yet also elucidate the core concepts so students can more efficiently solve the complex problems. Without intricate knowledge of student thinking, it is impossible to write subskill questions that are very fast and also very useful. In the failed implementations that we have seen, a lack of instructor knowledge in student difficulties and misconceptions led to lengthy subskill questions, a significant increase in the average time on task, and mixed reviews from students.

We have shown that the learning gains achieved by transforming homework using deliberate practice is equivalent to the gains from implementing active teaching strategies in class. Our data suggest that the introduction of subskills could, in some cases, result in a modest increase in the time on task on the order of 10%–15%. Notably, the time on task is similar even though students in the transformed homework group complete 3–4 times the number of problems as students in the traditional homework group. This illustrates that the time spent solving the subskill questions are a learning investment with immediate payoff in faster completion of the complex problems. Today, most college science, technology, engineering, and mathematics classes are using at least some active teaching strategies and the practice of applying the principles of deliberate practice to the in-class component is relatively common. It is time now to turn our attention to applying the same principles to the out-of-class component.

APPENDIX

The following are examples of complex problems and associated subskills encountered in this study. We discuss various implementations of subskills, including their features and potential failure modes.

1. Example of complex problem and relevant learning goals

The complex problem shown below, along with relevant learning goals, is taken from the Static Equilibrium module in Physical Sciences 2 (PS2).
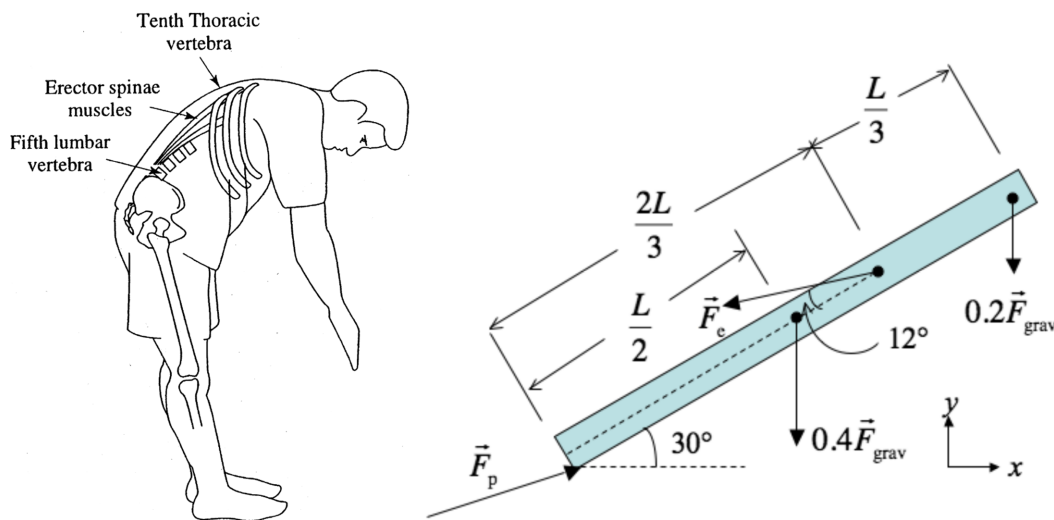

FIG. 5. Free-body diagram of the spinal column in a bend-forward position.

1. ***My aching back!*** (**3 pts**) When you bend over at the waist, large forces are exerted on the discs between your vertebrae. In particular, the force on the fifth lumbar vertebra can be quite substantial:

   The diagram above shows a free-body diagram for the spinal column. The force $\vec{F}_p$ is the force exerted by the pelvis on the base of the spine (at the fifth lumbar vertebra). Neither the direction nor the magnitude of this force is known. The force $\vec{F}_e$ is the force on the spine from the erector spinae muscles (the muscles that run along your back). The magnitude of this force is not known, but it is known to make an angle of 12° with the spine, as shown in the diagram. The other forces represent the fraction of your body weight supported by your spine; the vector $0.4\,\vec{F}_{grav}$ represents the weight of your trunk, while the vector $0.2\,\vec{F}_{grav}$ represents the weight of your head and arms. Assume a total body mass of 70 kg.

   a) Determine the magnitude of $\vec{F}_e$ while you maintain this position.
   b) Determine the magnitude and direction of $\vec{F}_p$.
   c) Explain, qualitatively, why holding a heavy object in your arms in this position is worse than holding the same heavy object in a back-pack on your back.

   The subset of learning goals relevant to the ***My aching back!*** complex problem is listed below:

   After this lecture, you will be able to…
   1. Identify if a system is in static equilibrium.
   2. Choose the most convenient coordinate system to analyze a rigid body in static equilibrium.
   3. Draw the free-body diagram of any rigid body and identify all the forces and their point of application, direction, and magnitude.
   4. Write all three equations* for static equilibrium and identify all knowns and unknowns.

Sub-skill: Can you find the most convenient choice for the pivot location?
Choosing a pivot location that allows one to solve for an unknown quantity in the fewest steps possible is a non-trivial task and is emphasized in the Static Equilibrium module of the course.

Question: A horizontal beam of length L and mass M = 50.0 kg is held in place by a pole and two wires, as shown below. Force $\mathbf{T_1}$ is at an angle of $\theta_1 = 45.0°$ and has an unknown magnitude. The only force you are given is $\mathbf{T_2}$, which is at angle of $\theta_2 = 30.0°$ and has a magnitude of 100. N.

Which choice of pivot location would allow you to solve for $\mathbf{T_1}$ using a single equation?

Location A
Location B
Location C
Location D
Location E
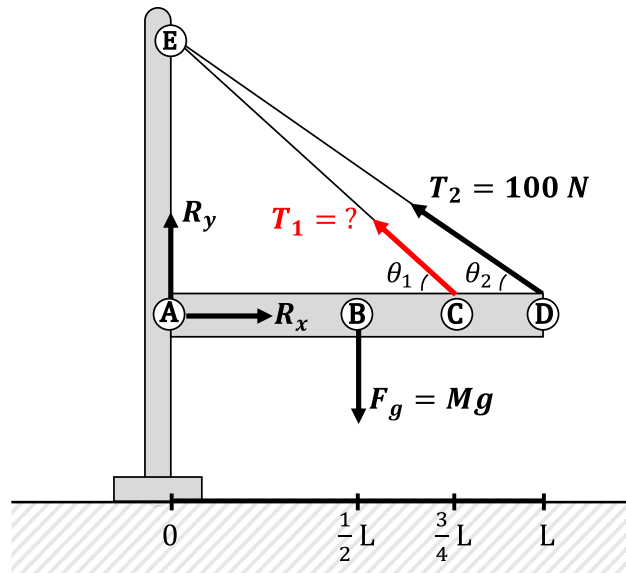
Hint 1 → Hint 2 → Hint 3   View solution

FIG. 6.   Example of a subskill targeting the *choice of pivot location*.

5. Identify the most convenient choice of pivot axis to use when applying $\Sigma\vec{\tau} = 0$ (this skill is non-trivial!).

*PS2 only studied static equilibrium in two dimensions. The three equations referenced here are $\Sigma F_x = 0$, $\Sigma F_y = 0$, and $\Sigma\vec{\tau} = 0$.

2. Subskill: Features and examples

The heading at the top of each subskill (in bold) is used to make explicit the connection between a subskill and its associated complex task(s) and application(s). In order to focus the student's attention on the cognitive task, the question stem is kept short, clear, and concise. Another critical part of subskills are the hints which provide students targeted feedback as they progress toward mastery. Each time a hint is revealed, ~5% is deducted from the point total. Each subskill contains anywhere from 1 to 4 hints, with the number depending on the relative difficulty and complexity of the problem. Writing hints that provide students with targeted and timely feedback requires expert knowledge of student thinking. This often requires the input from experienced teaching assistants and/or students from previous cohorts.

A detailed solution to the subskill problem is revealed once a correct answer is submitted, or when a student "gives up" and chooses to "view the solution," in which case a score of zero is assigned for the problem. As a guiding principle, each subskill problem, along with its hint(s) and solution, is written to ensure complete mastery (understanding) by the time students move on to the next subskill.

Two examples of subskills are shown in Fig. 6 and Fig. 7. They are both associated with learning goals practiced in the *My aching back!* complex problem.

Example 1 (See Fig. 6).

The subskill shown below targets learning goal #5 in the list above: "*Identify the most convenient choice of pivot axis to use for $\Sigma\vec{\tau} = 0$.*"

Hint 1: What forces (other than T_1) are missing information? What can you do to eliminate any need for this information?

Hint 2: Any force that is applied at the pivot location does not apply a torque and therefore does not show up in an equation that balances the torques about that point.

Sub-skill: Can you calculate torque?
Being able to calculate the torque of a force applied at an angle is an important skill when solving Rotation and Static Equilibrium-type problems.

Question: A cow is pushed to the right by a force F but does not move. Let the red dot indicated be the location for the pivot around which torque can be calculated.

Which expression correctly gives the magnitude of the torque due to F about the pivot?
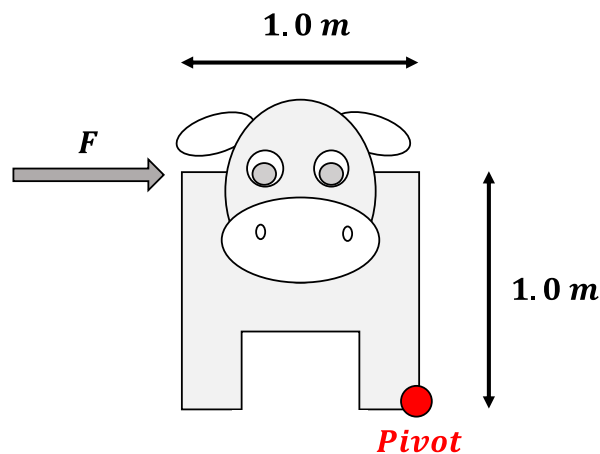
FIG. 7.   Example of a subskill targeting *torque with arbitrary moment arm.*

Where could you place the pivot so that some of the unknown forces are eliminated in this way?

Hint 3: Choose a pivot location so that R_x and R_y do not contribute any torque about that point. Then write an equation balancing torques and solve for T_1.

Solution: The answer is Location A. By choosing Location A as the pivot, the reaction forces R_x and R_y each have a torque of zero. If we then write an equation balancing the torques about Location A, we can easily solve for T_1:

$$\sum \vec{\tau} = 0$$

$$0 = -\tau_{\text{grav}} + \tau_{T_1} + \tau_{T_2}$$

$$0 = -Mg\left(\frac{L}{2}\right) + T_1\left(\frac{3L}{4}\right)\sin(180° - \theta_1)$$
$$+ T_2(L)\sin(180° - \theta_2)$$

$$0 = -(50.0\text{ kg})\left(9.81\frac{m}{s^2}\right)\left(\frac{1}{2}\right) + T_1\left(\frac{3}{4}\right)\sin(135°)$$
$$+ (100.N)(1)\sin(150°)$$

$$T_1 = 368\,N$$

Example 2 (See Fig. 7).
This next subskill targets learning objective #4: Write all three equations for static equilibrium and identify all knowns and unknowns.

Hint 1: What is the general equation for torque? How can you find the magnitude of a cross product?

Hint 2: Torque can be calculated as $\vec{\tau} = \vec{r} \times \vec{F}$. The magnitude of the torque can be calculated as $|\vec{\tau}| = |rF\sin\theta|$. What is r? What is θ?

Hint 3: The distance between the pivot and the point of application of the force is $r = \sqrt{2}$. The angle between $\vec{r}$ and $\vec{F}$ is 135°. Use these values in the equation for torque to find the answer.

Solution: The torque can be found as $|\vec{\tau}| = |\vec{r} \times \vec{F}| = |rF\sin\theta|$. The distance r between the pivot and where the force F is applied can be found using the Pythagorean theorem:

$$r = \sqrt{1^2 + 1^2} = \sqrt{2}.$$

The angle θ between the vectors $\vec{r}$ and $\vec{F}$ is $\theta = 90° + 45° = 135°$. Putting all of this together, we find

$$|\vec{\tau}| = |\vec{r} \times \vec{F}| = |rF\sin\theta| = \sqrt{2}F\sin(135°)$$
$$= \sqrt{2}F\left(\frac{\sqrt{2}}{2}\right) = \frac{2}{2}F$$
$$|\vec{\tau}| = F$$

3. Potential subskill failure modes

Occasionally, we have seen homework transformations that were rejected by students. From these

---

Sub-skill: Can you find which angles are equal to θ?
Using similar triangles to determine which angles are equal is an important skill needed for students to successfully use a tilted coordinate system when analyzing objects on an inclined plane.

Question: A block rests on an incline as shown. The incline makes an angle θ with the horizontal as shown. Two triangles have been drawn using dashed lines.

Select all angles that are also equal to θ.

☐ A
☐ B
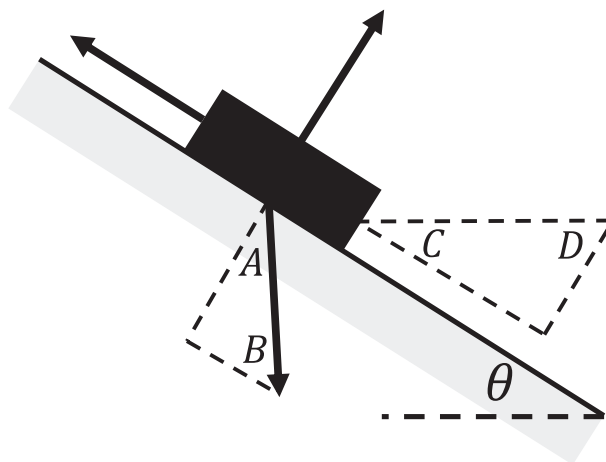☐ C
☐ D
☐ None of these

FIG. 8.   Example of a failed subskill with a difficulty level ill-suited for a specific student population.

Sub-skill: Can you calculate potential energy?
This is the first part in a multi-step tutorial problem for using conservation of energy to solve ultimately for the mass of the ball.

Question: A ball of unknown mass m is tossed straight up with initial speed v. At the moment it is released, the ball is a height h above a spring-mounted platform. The ball rises, and falls back toward the platform, ultimately compressing the spring a maximum distance d from its relaxed position. Assume that the spring is perfectly ideal with spring constant k, and that the mass of the spring and platform is negligible.

Consider the system *at the moment the ball is released*, as illustrated in the figure at right. At this instant, the system has multiple forms of energy. Find expressions for the following forms of energy in terms of m, g, d, h, v, and k.

What is the gravitational potential energy $E_{g,i}$ of the system at the moment the ball is released? Assume that the *zero* gravitational energy corresponds to the ball being at the level of the platform in its relaxed state.
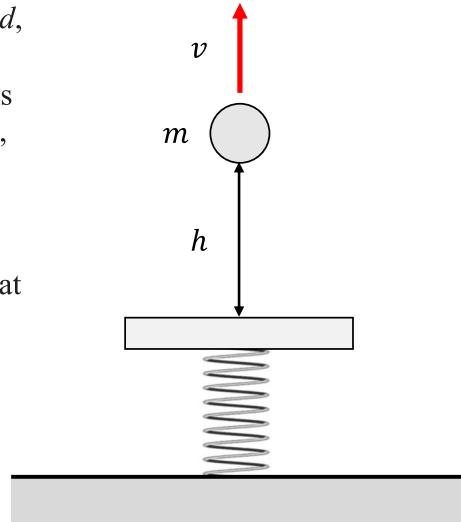
$$E_{g,i} = \boxed{}$$

FIG. 9.   Example of a failed subskill with an overwhelming amount of text and details.

Sub-skill: Can you apply conservation of momentum and the work-energy theorem?
This problem contains two important concepts that appear repeatedly in the course: conservation of momentum and the work-energy theorem.

Question: Block 1 has a mass of 1.0 kg and slides at a constant speed of 1.0 m/s along ice. It then collides with Block 2, which is initially stationary and has a mass of 2.0 kg. The two blocks stick together and then move across a rough patch with $\mu_k = 0.50$ until they come to rest.

What is the speed of the two blocks directly after the collision?

1.0 m/s

$\boxed{}$ m/s
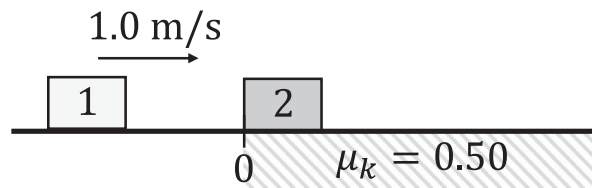
How far do the blocks travel before they come to rest?

$\boxed{}$ m

FIG. 10.   Example of a failed subskill that attempts to target multiple skills simultaneously.

implementations, we have identified several mechanisms through which a subskill can fail.

For each potential failure mode listed below, we have included an example and commentary. The examples are taken from a failed homework transformation in the course Physics 15a, an advanced introductory mechanics class for physics majors at Harvard University. For perspective, Harvard offers several other introductory physics courses aimed at different student populations: PS2 and PS3 is a year-long calculus-based introductory physics sequence for biological sciences majors; E1ax is an algebra-based introductory mechanics course offered to online students; and Physics 12a is aimed at engineering students.

Failure mode 1 (See Fig. 8): Mismatch between student population and subskill difficulty. In such cases, the subskill may be too easy (and therefore unproductive) or too difficult for a given student population.

The difficulty of this subskill is more consistent with the level of expertise typically demonstrated by students in E1ax or PS2. Students in these courses often have trouble identifying equal angles for objects when using a tilted coordinate system; for the advanced students in 15a, however, this question is trivial and unproductive. This particular example shows how the same question can be great subskill for more novice students and yet a terrible subskill for more advanced students. Having intimate knowledge of common misconceptions and difficulties for a student population is key when writing effective subskills for them.

Failure mode 2 (See Fig. 9): The interpretation and parsing of text within the subskill requires effort. This increases the cognitive load and distracts from the subskill itself.

This straightforward subskill, which asks students to calculate the gravitational potential energy of a ball at a certain height h, is obscured by the overwhelming amount of text. Students must the carefully read and interpret more than 150 words before realizing what the question is asking. Problems that involve dense blocks of text should be avoided in favor of direct, simple-to-interpret questions.

Failure mode 3 (See Fig. 10): The subskill targets multiple skills and does not provide feedback on one specific skill.

In this example subskill, two separate skills are targeted: conservation of momentum and the work-energy theorem. This question essentially contains two subskills within a single problem. A better alternative would be to ask each part of this question separately.

Failure mode 4: Including more subskills beyond the "essential" 25–30 needed for a module. The likely increase in the time on task can result in students being discouraged and frustrated, and may only provide marginal benefits.

Any subskill that is included beyond the recommended limit can serve as an example for this failure mode. For example, a transformed homework comprised of 50 perfectly written subskills will often lead to worse outcomes than one consisting of the recommended 25–30 problems. Good subskills provide targeted feedback, so students should be able to master the tasks after a few questions and therefore do not need the additional problems that add little benefit to their understanding and increase the time on task.

[1] R. R. Hake, Interactive-engagement vs. traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses, Am. J. Phys. **66**, 64 (1998).

[2] C. H. Crouch and E. Mazur, Peer Instruction: Ten years of experience and results, Am. J. Phys. **69**, 970 (2001).

[3] L. Deslauriers, E. Schelew, and C. Wieman, Improved learning in a large-enrollment physics class, Science **332**, 862 (2011).

[4] S. Freeman, S. Eddy, M. McDonough, M. Smith, N. Okoroafor, H. Jordt, and M. Wenderoth, Active learning increases student performance in science, engineering, and mathematics, Proc. Natl. Acad. Sci. U.S.A. **111**, 8410 (2014).

[5] J. M. Fraser, A. L. Timan, K. Miller, J. E. Dowd, L. Tucker, and E. Mazur, Teaching and physics education research: bridging the gap, Rep. Prog. Phys. **77**, 032401 (2014).

[6] L. Deslauriers, L. S. McCarty, K. Miller, K. Callaghan, and G. Kestin, Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom, Proc. Natl. Acad. Sci. U.S.A. **116**, 19251 (2019).

[7] K. A. Ericsson, R. T. Krampe, and C. Tesch-Römer, The role of deliberate practice in the acquisition of expert performance, Psychol. Rev. **100**, 363 (1993).

[8] K. A. Ericsson, Deliberate practice, and acquisition of expert performance: A general overview, Academic Emergency Med. **15**, 988 (1998).

[9] C. Wieman, Applying new research to improve science education, Issues Sci. Technol. **29**, 25 (2012).

[10] C. E. Wieman, Expertise in university teaching & the implications for teaching effectiveness, Eval. Training Daedalus **148**, 47 (2019).

[11] L. Deslauriers and C. Wieman, Learning and retention of quantum concepts with different teaching methods, Phys. Rev. ST Phys. Educ. Res. **7**, 010101 (2011).

[12] J. D. Jones, W. K. Madison, and C. Wieman, Transforming a fourth-year modern optics course using a deliberate

practice framework, Phys. Rev. ST Phys. Educ. Res. **11,** 020108 (2015).

[13] B. D. Mikula and A. F. Heckler, Framework and implementation for improving physics essential skills via computer-based practice: Vector math, Phys. Rev. Phys. Educ. Res. **13,** 010122 (2017).

[14] L. L. Parker and G. M. Loudon, Case study using online homework in undergraduate organic chemistry: Results and student attitudes, J. Chem. Educ. **90,** 37 (2013).

[15] G. Kortemeyer, E. Kashy, W. Benenson, and W. Bauer, Experiences using the open-source learning content management and assessment system LON-CAPA in introductory physics courses, Am. J. Phys. **76,** 438 (2008).

[16] E. Morote and D. E. Pritchard, Technology closes the gap between students' individual skills and background differences, in *Proceedings of the Society for Information Technology & Teacher Education International Conference* (Association for the Advancement of Computing in Education, Chesapeake, 2004), pp. 826–831.

[17] K. Vanlehn, C. Lynch, K. Schulze, J. A. Shapiro, R. Shelby, L. Taylor, D. Treacy, A. Weinstein, and M. Wintersgill, The Andes physics tutoring system: Lessons learned, Int. J. Artif. Intell. Educ. **15,** 147 (2005).

[18] B. Bloom, Learning for mastery, Eval. Comment **1,** 1 (1968), https://files.eric.ed.gov/fulltext/ED053419.pdf.

[19] https://www.macmillanlearning.com/college/us/digital/sapling.

[20] D. Hestenes, M. Wells, and G. Swackhamer, Force Concept Inventory, Phys. Teach. **30,** 141 (1992).

[21] G. Kestin, K. Miller, L. S. McCarty, K. Callaghan, and L. Deslauriers, Comparing the effectiveness of online versus live lecture demonstrations, Phys. Rev. Phys. Educ. Res. **16,** 013101 (2020).

[22] L. S. McCarty and L. Deslauriers, Transforming a large university physics course to student-centered learning, without sacrificing content, in *The Routledge International Handbook of Student-Centered Learning and Teaching in Higher Education* (Routledge, London, 2020), pp. 186–200.

[23] L. Ding, R. Chabay, B. Sherwood, and R. Beichner, Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment, Phys. Rev. ST Phys. Educ. Res. **2,** 010105 (2006).

[24] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevPhysEducRes.17.010129 for the tests of learning used in this study along with an example of an interactive/transformed lecture.

[25] W. K. Adams, K. K. Perkins, N. S. Podolefsky, M. Dubson, N. D. Finkelstein, and C. E. Wieman, New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey, Phys. Rev. ST Phys. Educ. Res. **2,** 010101 (2006).

[26] M. Richards-Babb *et al.* Online homework, help or hindrance? What students think and how they perform, J. Coll. Sci. Teach. **40,** 81 (2011).

[27] M. Mendicino, L. Razzaq, and N. T. Heffernan, A comparison of traditional homework to computer-supported homework, J. Res. Tech. Educ. **41,** 331 (2009).

[28] K. Kelvin Cheng, B. Ann Thacker, R. L. Cardenas, and C. Crouch Using an online homework system enhances students' learning of physics concepts in an introductory physics course, Am. J. Phys. **72,** 1447 (2004).

[29] W. K. Adams and C. E. Wieman, Analyzing the many skills involved in solving complex physics problems, Am. J. Phys. **83,** 459 (2015).

[30] K. A. Ericsson, The influence of experience and deliberate practice on the development of superior expert performance, *The Cambridge Handbook of Expertise and Expert Performance (Cambridge Handbooks in Psychology)*, edited by K. A. Ericsson, N. Charness, P. J. Feltovich, and R. R. Hoffman (Cambridge University Press, Cambridge, England, 2006), pp. 683–703, https://doi.org/10.1017/CBO9780511816796.038.

[31] K. A. Ericsson, Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains, Acad. Med. **79,** S70 (2004).

[32] M. R. Lepper and M. Woolverton, The wisdom of practice: Lessons learned from the study of highly effective tutors, in *Improving Academic Achievement*, edited by J. Aronson (Academic Press, New York, 2002), pp. 135–158.

[33] B. W. Wood and K. D. Tanner, The role of the lecturer as tutor: Doing what effective tutors do in a large lecture class, CBE Life Sci. Educ. **11,** 3 (2012).