

Physics exam preparation: A comparison of three methods

Witat Fakcharoenphol and Timothy Stelzer

Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA

(Received 15 August 2013; published 17 March 2014)

In this clinical study on helping students prepare for an exam, we compared three different treatments. All students were asked to take a practice exam. One group was then given worked-out solutions for that exam, another group was given the solutions and targeted exercises to do as homework based on the result of their practice exam, and the third group was given the solutions, homework, and also an hour of one-on-one tutoring. Participants from all three conditions significantly outperformed the control group on the midterm exam. However, participants that had one-on-one tutoring did not outperform the other two participant groups.

DOI: [10.1103/PhysRevSTPER.10.010108](https://doi.org/10.1103/PhysRevSTPER.10.010108)

PACS numbers: 01.40.Fk, 01.50.H-, 01.40.gb

I. BACKGROUND

Practicing old exam problems can be an effective activity in preparing for an exam. From a class survey, 93% of students in Introductory Physics at the University of Illinois at Urbana-Champaign rate practice exams as useful to essential in helping them prepare for an exam. The survey agrees with Ericsson *et al.* [1] that becoming an expert or mastering a skill requires deliberate practice. To practice on old exam problems effectively, students need to treat the practice problems as if they were an actual exam. By doing so, students could benefit from the testing effect [2] by identifying knowledge gaps and transferring knowledge to new situations. Other benefits from the testing effect include improving retention, knowledge organization, and future learning. Mistakes made in the practice exams do not penalize students in any way. Rather, it provides students the formative assessment that could help them focus on increasing competence. As a result, taking practice exams encourages students to be mastery oriented [3] rather than showing helpless behavior [4]. Thus, deliberate practice on old exam problems by focusing on improving competence and not worrying about making mistakes could improve students' performance on the exam.

In addition to the practice gained by doing old exam problems, feedback on the students' work is also important for learning. Indeed, Epstein *et al.* [5] found that practice without feedback does not improve performance. For students in the introductory physics classes at University of Illinois at Urbana-Champaign, practice exams usually come with an answer key, which provides a minimum level of feedback—only indicating the correct answer. A richer form of feedback might improve the learning outcomes.

Since worked examples have been shown to be an effective learning tool [6,7], we created a web-based activity that provides a set of organized old exam problems, along with their worked solutions, for students to study.

This practice exam system significantly improved students' performance on closely related problems [8]. However, the immediate gains from the practice problems and worked-out solutions were limited. Students improved significantly on problems with identical solution steps, whereas they improved relatively little on problems with nonidentical solution steps. In order to help students generalize the techniques they learn in the practice exam questions, a better form of feedback might be needed.

One form of feedback that could be more effective than worked-out solutions is the use of one-on-one tutors. Tutors can check students' work on the practice exam, assess students' knowledge, and customize their feedback and guidance to suit what students need. Wood, Bruner, and Ross [9] also suggest that tutoring can help break down a complex problem to fit the tutee's ability, emphasizing critical features and demonstrating how to solve the problem. Although private tutoring may not be a practical solution for an entire class, it provides an important benchmark for measuring the effectiveness of other, more practical options.

Individualized feedback is an option suitable for a large class with online capability. Since testing (practice exams in this case) can identify students' knowledge gaps [2] and students can improve by doing practice exams [8], then more practice, individually tailored to the topics each student struggles with, might be another useful form of feedback. It can also be implemented using web-based practice exams. This paper describes the results of a clinical study designed to test the impact of these treatments on student exam performance.

Unlike previous physics education research (PER) studies designed to improve problem-solving skills through different styles of problems [10,11], feedback [12], and class dynamics [11,13,14], our intervention focused on

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

helping students prepare for the exam, after they had already completed all of the course-required learning activities. This research builds on the testing effect [3] and learning activities such as worked examples [6,7] and tutoring [9]. In particular, it investigates if taking practice exams in an examlike environment and providing full solutions as feedback significantly improves student performance on the actual exam. It also provides data on the relative impact of providing students with additional learning resources including customized homework (HW) problems and one-on-one tutoring. Analysis of student interactions with and performance on the learning resources provides additional information about the effectiveness and limitations of the various activities.

II. EXPERIMENT

A. Setting

The experiment setting is summarized in Fig. 1.

For this experiment, we compared three different sets of exam preparation activities based on practice exams: (1) practice exams with solution feedback (practice only), (2) practice exams with solution feedback and related homework (practice plus homework), and (3) practice exams with solution feedback, related homework, and tutor sessions with an experienced PER member (practice plus homework plus tutor). The experiment was a clinical study meeting three times in the week before the second midterm. There were three sets of practice exam problems. They were all multiple-choice problems similar to those given on past exams. All three practice exams were considered equivalent in difficulty and coverage. The solution feedback contained the basic strategy to solve the problems and

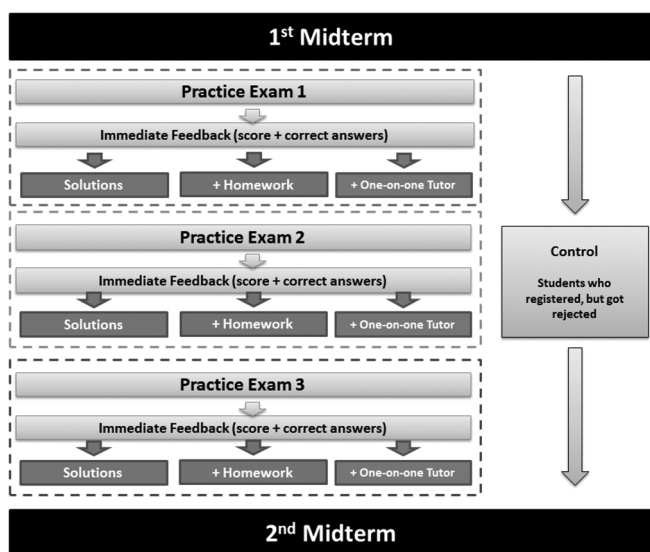


FIG. 1. Experiment setting. The sequence of the activities of each group flowed from top to bottom. Participants in each group received one type of treatment throughout the experiment.

a couple of steps and equations to get to the final answer. The related homework feedback consisted of similar exam problems with the final answers already given. The participants could work on the problems and check the answers by themselves after the experiment sessions. Participants with the tutor condition had a one hour session with a tutor after every practice exam to go over the practice problems. The answers and confidence level on all of the practice problems were collected. The score on the following midterm exams of all participants was analyzed.

B. Population

The experiment was done in the calculus-based Introduction to Classical Mechanics course at the University of Illinois at Urbana-Champaign. This course is required for both physics and engineering majors, typically taken in their first year at the university. In Fall 2012, students who scored less than the average score in the first midterm were invited by Email to participate in the experiment. The invitation was sent out two weeks before the second midterm. Seventy-six students [black bars in Fig. 2(a)] were randomly assigned into four groups: (1) practice exam with solution feedback, (2) practice exam with solution feedback and related homework, (3) practice exam with solution feedback, related homework, and

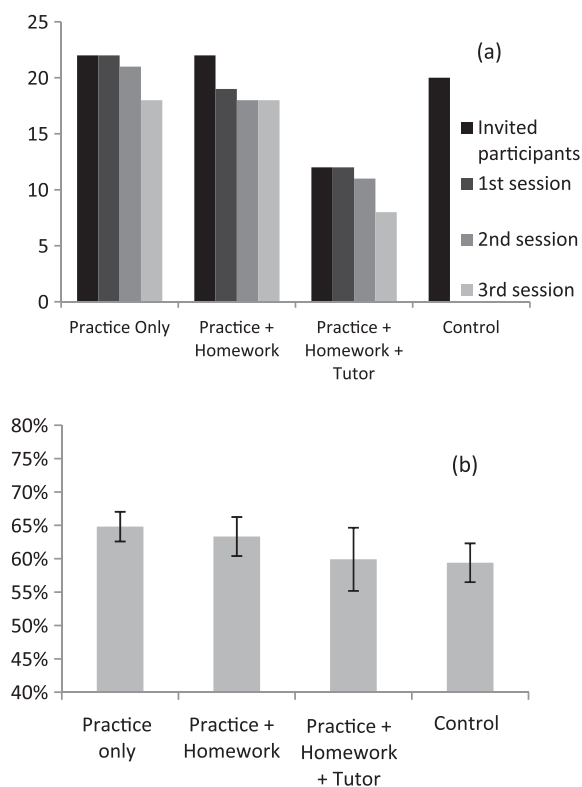


FIG. 2. (a) Number of participants for each group. Participants who missed the previous session were not allowed to participate in the rest of the experiment. (b) Initial score on the first midterm (incomplete participants included).

tutoring sessions, and (4) no treatment as the control group. The number of participants for each group is shown in Fig. 2(a). Three participants who did not show up at the first session and two participants who dropped the class (one from the practice only group and one from the tutor group) were discarded from the analysis. Seven participants who did not complete all of the tasks were also excluded. The average scores on the first midterm of the 44 participants who completed all three treatment sessions [yellow bars in Fig. 2(a)] and 20 participants in the control group are shown in the graph in Fig. 2(b).

The students in the first three groups received an Email telling them that they had been selected to participate in the experiment and that they would receive compensation if they participated in all three experiment sessions. We chose to tell students about the compensation after they were selected to filter out students who might participate in the experiment only for the money. The students in the control group received an Email explaining the experiment was full and they were not able to participate in any activities of the experiment. Randomly selecting registered students as the control minimized the self-selection effect. Note that the exams for the past ten semesters and the key were available to all students in the course.

C. Procedure

The participants attended three sessions on three different days. At the beginning of each session, they answered some survey questions about their exam preparation. Then they worked on the practice exam for one hour in an examlike environment. For each problem, they had to choose an answer and their confidence in their answer. We used a set of three practice exams, one for each session. The order of the practice exams was random for each student. After one hour, the practice exams were graded and given back immediately along with the worked-out solutions. Then the participants that did not get tutoring left for the day.

Participants in the practice plus homework group received a maximum of three related homework problems in addition to the worked-out solution. These related homework problems were selected based on the participants' answers and confidence. The homework related to the problems that the students got wrong but had high confidence were given out first. These homework problems are similar to the problems on the practice exams. The final answers, but not worked-out solutions, were also provided with the homework problem text. Participants were expected to show their work on the homework and return it at the next experiment session.

Each participant in the practice plus homework plus tutor group spent an hour with a tutor immediately after working on the practice exam. Over three sessions, each participant received three hours of individual tutoring. The participant could also work on the practice problems with the tutor or ask the tutor physics questions. At the end of the session,

the tutor gave the participant a maximum of three related homework problems based on the tutor's judgment.

After three experiment sessions, all of the participants and the control group took the second midterm. All of the problems on the midterm covered the same topics as the practice exams, but they were not similar to any problems used in the experiment.

D. Results

The students' scores on the three practice exams showed gradual improvement [Fig. 3(a)]. All three participant groups scored higher from session to session, and there was no significant difference between the groups. This is consistent with our earlier findings, since the problems on the three versions of the exam were designed to have similar solutions. If we consider the scores from the first practice exam and the third practice exam, all three participant groups showed significant gain [Fig. 3(b)] [$t_{\text{practice only}}(\text{DOF} = 17) = 3.4$, $p < 0.01$], [$t_{\text{practice+HW}}(\text{DOF} = 17) = 4.6$, $p < 0.001$], [$t_{\text{tutor}}(\text{DOF} = 7) = 3.9$, $p < 0.01$].

On the actual midterm exam, the participants performed significantly better than the control. On average, the participants showed a positive change from the first midterm to the second midterm. In contrast, students in the control group showed a drop [Fig. 4(a)]. On average, the participants scored $3.2 \pm 1.5\%$ higher and the control scored $2.8 \pm 2.5\%$ lower from the first to the second

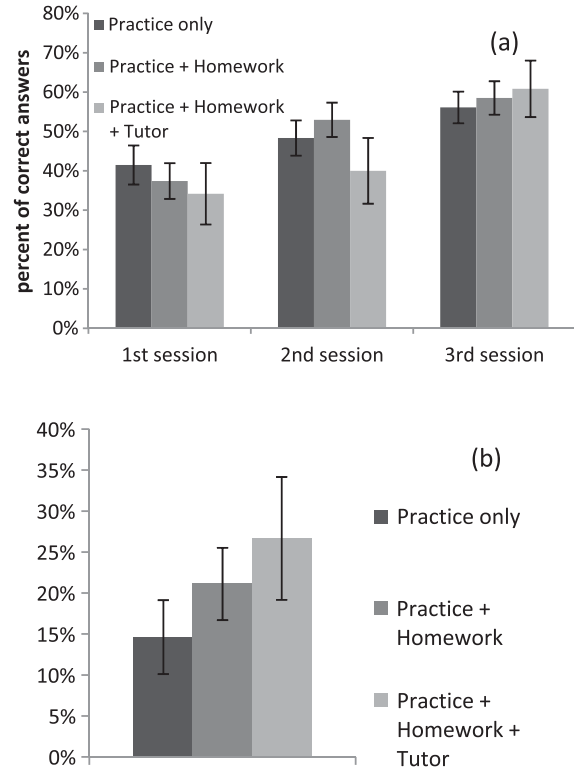


FIG. 3. (a) Practice exam scores. (b) Practice exam score changes between the first and third sessions.

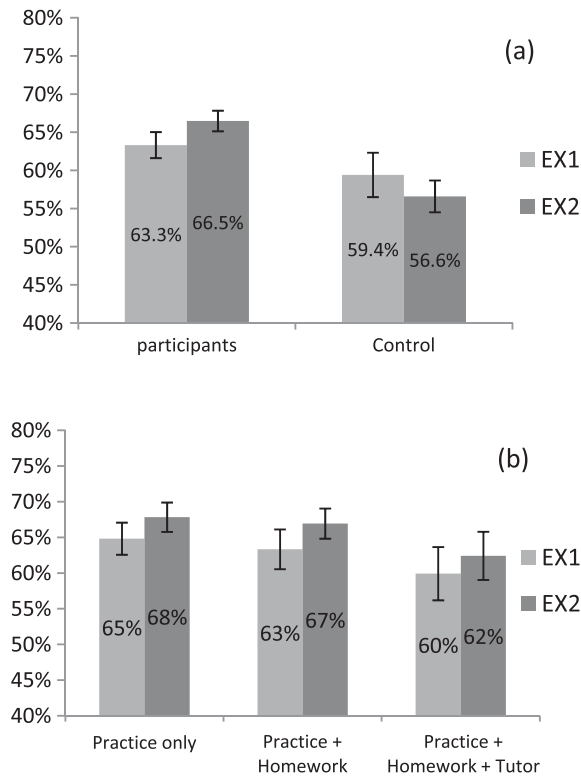


FIG. 4. (a) First and second midterm exam scores (before and after the treatment) from participants who completed three treatment sessions along with the control group. (b) First and second midterm from three treated groups.

midterm. The score drop of the control group was consistent with the average score drop of the whole class ($-5.9 \pm 0.4\%$). It is important to note that none of the midterm problems had solution steps identical to any of the problems provided in the treatment sessions. By comparing the participants to the control group, we conclude that the treatments significantly helped students improve their exam performance [$t(\text{DOF} = 62) = 2.85, p < 0.01$].

However, participants with a personal tutor did not outperform the other two participant groups [Fig. 4(b)]. Because of a 30% dropout rate of the tutor treatment group and the fact that the participants in the control group could not quit the experiment, this self-selection effect cannot be ignored. When we reanalyze the data to include all invited participants (except two participants who dropped the class), the tutor group scored significantly less than the practice plus homework group [$t(\text{DOF} = 31) = 2.32, p < 0.05$]. Note that participants who completed three sessions scored about 6% higher [$t(\text{DOF} = 62) = 2.85, p < 0.01$] compared to the control group [Fig. 4(a)]. When we included all invited participants, the gains of the three participant groups was reduced to $4.4 \pm 2.1\%$ compared to the control group [$t(\text{DOF} = 72) = 2.15, p < 0.05$].

The data presented thus far show that students in the treatment group outperform those in the control group. This

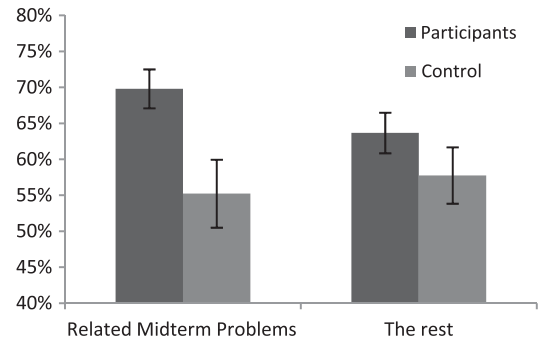


FIG. 5. Performance on the second midterm categorized by similarity to the practice exam.

finding was an extension to the work of Fakcharoenphol *et al.* [8]. It showed that practice problems can lead to improvement on the midterm exam problems that did not have solutions identical to the practice problems. However, analysis based on the total exam score does not reveal any significant differences between the three treatments. In order to better discern the effect of the different treatments, the problems given on the midterm were divided into two groups, based on if the key concepts necessary for solving that problem were covered in the practice test or not. Two judges independently categorized the problems with roughly half of the problems being placed in each category. The agreement rate between two judges was initially 79% and the disagreement was discussed and resolved. Figure 5 shows that the performance gain of the treatment group relative to the control group was about twice as large on the related problems (15%) compared to the unrelated problems (8%). The following analysis will focus on the results of the ten midterm exam problems categorized as related to the treatment activities, in order to better understand the impact of the different treatments.

Figure 6 shows the impact that targeted homework activities had on student exam performance. The first two bars show that for students that got a particular topic wrong, those that completed a targeted homework problem on that topic scored 10% higher than those who were not given a homework assignment targeting that topic. For each problem, we categorize participants into three groups: those

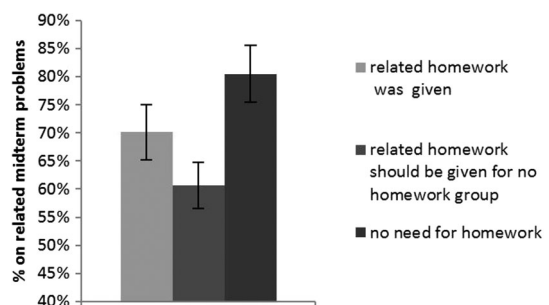


FIG. 6. Performance on related midterm problems.

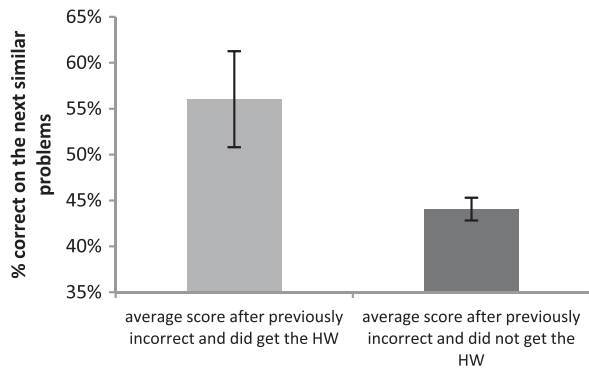


FIG. 7. Performance on later practice problems if previously incorrect.

who needed and received related homework, those who needed but did not receive the homework, and those who did not need the homework. We assume that participants who answer incorrectly on the practice problem need the corresponding related homework and participants who answer correctly on the practice problem do not need that related homework. We find that participants who did not need the related homework scored the highest on average across all related midterm problems. Participants who needed and received the homework scored significantly higher than participants who needed the homework but did not receive it [$t(\text{DOF} = 9) = 2.99, p < 0.05$].

The impact of the related homework activities can also be observed in the student performance on the practice exams during the experiment (Fig. 7). Looking at each case where a student got a question wrong on the practice test ($N = 445$), participants who received related homework during the previous session performed about 10% better on the targeted problems on the next session than participants who did not get the homework [$t(\text{DOF} = 443) = 2.05, p < 0.05$]. We conclude that providing targeted homework activities can significantly improve student performance on that topic.

E. ONE-ON-ONE TUTORING RESULTS

Despite the fact that participants in the tutoring group spent more time one on one with a tutor and received the same set of related homework, they did not perform better than the other two treatment groups. Participants in the tutoring group spent three extra hours, one hour at the end of every session, working one on one with tutors. The tutors in this experiment were members of the physics education research group at University of Illinois at Urbana-Champaign. In the session, the tutors answered participants' physics questions and explained physics concepts. They helped participants work through the practice problems. They also assigned related homework problems, used in the practice plus homework group, and checked them with the tutee in the subsequent experiment sessions. It is

an interesting question as to why the students that were tutored did not perform better.

One explanation is that the tutors did not do a good job. Although this may be true, it is a somewhat unsatisfactory conclusion, as the tutors were all highly trained in physics education and were engaging in best practices based on the student responses. The tutors not only knew the solution to all of the practice problems, they also discussed the common misconceptions and mistakes before every session. Within the session, tutors encouraged students to show their work and explain their reasoning. Students were encouraged to ask if they did not understand any part of the material. The tutors also asked the students to do similar problems to gauge their understanding. At the end of the session, the tutors gave their students homework problems targeting each student's weak topics. Anecdotally, the tutors commented that during the sessions many of the students were unable to concentrate due to lack of sleep or concern about completing work for other courses. This observation is consistent with the relatively high attrition rate of 30% for the tutor treatment group. Survey questions administered before each session indicate that students from all three treatment groups spent about the same amount of time preparing for the exam. The total time spent between the first and third sessions is shown in Fig. 8. All three groups spent, on average, 10 hours total. Related homework and tutoring time did not significantly change the total time participants spent preparing for the exam. However, the participants in groups with related homework (practice exam, solution, and homework) spent more time on materials provided from the experiment (bottom section of the time bars) than participants without homework [$t(\text{DOF} = 34) = 2.8, p < 0.01$]. Also, participants in the tutor group spent significantly less time beyond the experiment material than the other two treatment groups (top section of the time bars) [$t(\text{DOF} = 42) = 2.25, p < 0.05$]. Although the three groups spent about the same total time preparing for the exam, they distributed the time on activities differently.

It is difficult to draw strong conclusions about the impact of three hours of tutoring based on the data collected in this

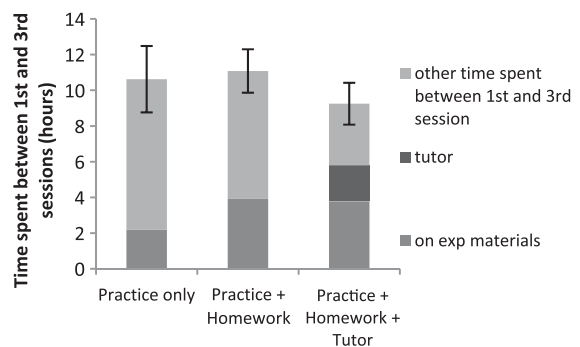


FIG. 8. Self-report on time spent between the first and third sessions.

experiment. However, the data do suggest two factors that should be considered in designing an effective tutoring treatment. First, the high attrition rate and relatively poor condition of students attending the sessions suggest that student time constraints seem to be a very important factor in their learning. Second, the large reduction in time spent studying outside the course suggests students being tutored may be overly reliant on the tutor to ensure they understand the material, instead of taking responsibility themselves.

III. DISCUSSION

The idea of providing more practice opportunities and learning tools on the most needed topics is supported by the result of this experiment. As we expected from our previous work, participants perform better on the subsequent experiment session on similar problems with identical solution steps. However, in this experiment, we also see an improvement on the related midterm problems, which have non-identical solution steps to the practice problems. Related homework exercises, specifically selected for each student, improved the students' performance on related midterm problems.

One-on-one tutoring, on the other hand, does not guarantee an extra boost in this practice exam scenario. Despite the extra time the participants spent with tutors whom we considered more experienced than average tutors and the extra time they used to work on the related homework, these participants performed only as well as the practice only group. The difference in "other time" suggests that tutoring or the idea of having a tutor can change the way these students prepared for the exam or at least the way they self-report their time.

In a previous experiment [8], a web-based exam preparation tool was available to all students but its use was optional. In that study the self-selection effect would bias any results comparing later midterm scores of the users and nonusers of the practice exams. The current experiment was conducted in a clinical condition to minimize the self-selection effect. All treated students and the control group were randomly selected from the same group of students who registered for the experiment. As a result, the effect of the treatment on midterm exam performance could be measured and compared.

The goal of this experiment is to increase the performance on the later problems with nonidentical solution steps [8]. No midterm problems had identical solution steps to any practice problems in this experiment. The midterm and practice problems covered the same physics concepts, but no solutions given during the experiment can be directly used to solve any of the midterm problems. The examples of problems with identical solution steps are the three sets of practice problems and the targeted homework problems (see example problems in the Supplemental Material [15]).

A few important differences from our previous experiment are that the practice exams were given in an examlike environment three times before the midterm instead of available online a week before the midterm. Student performance on the midterm was measured and compared in this experiment, whereas the performance on the next similar practice problems was used in the previous experiment. Also, this current experiment was in an introductory mechanics course instead of the introductory electromagnetism course.

There are many possible reasons why the participants with tutoring did not outperform other participants on the exam. One thing that can go wrong in a one-on-one tutoring session is tutees' false inferring of what the students understand. Person *et al.* [16] found that the quality of students' questions and the students' answer to the comprehension-gauging questions (e.g., Do you understand?) can be misleading in regards to their real understanding. For this experiment, tutors had only one hour for each session to gauge their tutee's skill, go over all the practice problems, and suggest what the tutees should study more. Rushing through these processes can cause errors in each tutoring step and result in ineffective tutoring.

The time for tutoring might be another important key. In the experiment, the tutoring session is right after an hour of practice exam in an examlike environment. By the time they come to the tutoring session, participants may have already used up most of their problem-solving resources. Learning when they were already exhausted might not be as useful as it should be.

Another possible problem is due to the tutees' idea of having a tutor. Since they have spent hours with tutors, or they know that they will have tutors again and they will get help for sure, they might not study by themselves as hard as they could. The self-report shows that the time spent on material beyond the experiment (Fig. 8) for the group with tutors is significantly less than other participant groups.

The tutors in this experiment are all experienced tutors and can be considered above-average tutors in normal classrooms. Also, spending three hours of tutoring per student is costly from the instructor's point of view. For these reasons and from the results of the experiment, one-on-one tutoring might not be the gold standard for physics learning at this stage.

Similar to the idea of having a tutor, the idea of rejection from a possible source of help can also be a factor that hindered the performance of the control group. We do not have evidence to support or reject this idea, except that the score drop of the control group is in the same trend as the score drop of the whole class.

Beyond the resources provided in this experiment, students normally have access to the past midterm problems of the past ten semesters. Every week, they also have a normal discussion session where they work in a group solving problems. They also have office hours in which

they can ask questions and get help from the teaching assistants.

IV. FUTURE IMPROVEMENT

Time availability is a major constraint for students. The future practice exam system should be “efficient” and available for students at a time convenient for them. Right now, old exam problems and the key are available to students, but they are organized by the semesters they were created, which are not by topic. If students struggle through a problem in a set and believe that they are capable of solving it, they have to continue working on different sets of problems before they can find a similar problem to test their understanding. Also, the key given with the old

exam problems could contain more guidance information beyond the correct answer.

An adaptive practice exam system might be the next important step. A difficulty-gauging system could help students get to a more suitable starting point, for their skill level, faster. High-performance students can start at more challenging problems, and low-performance students can practice more on fundamental conceptual problems. This might help to motivate students to use the practice problems more, especially for what they need.

ACKNOWLEDGMENTS

This work was supported by the NSF through Grant No. NSF DUE 08-17185.

-
- [1] K. Ericsson, R. Krampe, and C. Tesch-Römer, The role of deliberate practice in the acquisition of expert performance, *Psychol. Rev.* **100**, 363 (1993).
 - [2] H. Roediger III, A. Putnam, and M. Smith, Ten benefits of testing and their applications to educational practice, *Psychol. Learn. Motiv.* **55**, 1 (2011).
 - [3] E. Elliott and C. Dweck, Goals: An approach to motivation and achievement, *J. Pers. Soc. Psychol.* **54**, 5 (1988).
 - [4] C. Diener and C. Dweck, An analysis of learned helplessness: Continuous changes in performance, strategy, and achievement cognitions following failure, *J. Pers. Soc. Psychol.* **36**, 451 (1978).
 - [5] M. Epstein, A. Lazarus, T. Calvano, K. Matthews, R. Hendel, B. Epstein, and G. Brosvic, Immediate feedback assessment technique promotes learning and corrects inaccurate first responses, *Psychol. Rec.* **52**, 187 (2002).
 - [6] F. Paas and T. Van Gog, Optimising worked example instruction: Different ways to increase germane cognitive load, *Learn. Instr.* **16**, 87 (2006).
 - [7] T. Van Gog, F. Paas, and J. Vanmerrienboer, Effects of process-oriented worked examples on troubleshooting transfer performance, *Learn. Instr.* **16**, 154 (2006).
 - [8] W. Fakcharoenphol, E. Potter, and T. Stelzer, What students learn when studying physics practice exam problems, *Phys. Rev. ST Phys. Educ. Res.* **7**, 010107 (2011).
 - [9] D. Wood, J. Bruner, and G. Ross, The role of tutoring in problem solving, *J. Child Psychol. Psychiatry* **17**, 89 (1976).
 - [10] D. Maloney, C. Hieggelke, and S. Kanim, nTIPERS: Tasks to help students “unpack” aspects of Newtonian mechanics, *AIP Conf. Proc.* **1289**, 33 (2010).
 - [11] P. Heller and M. Hollabaugh, Teaching problem solving through cooperative grouping. Part 2: Designing problems and structuring groups, *Am. J. Phys.* **60**, 637 (1992).
 - [12] To experience the interactive examples, please visit <http://research.physics.illinois.edu/per/ie.html>.
 - [13] P. Heller, R. Keith, and S. Anderson, Teaching problem solving through cooperative grouping. Part 1: Group versus individual problem solving, *Am. J. Phys.* **60**, 627 (1992).
 - [14] M. Scott, T. Stelzer, and G. Gladding, Explicit reflection in an introductory physics course, *AIP Conf. Proc.* **951**, 188 (2007).
 - [15] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevSTPER.10.010108> for examples of problems used in the experiment: three practice problems with identical solutions steps, a targeted homework, and a related midterm problem.
 - [16] N. Person, A. Graesser, J. Magliano, and R. Kreuz, Inferring what the student knows in one-to-one tutoring: The role of student questions and answers, *Learning and Individual Differences* **6**, 205 (1994).