

## Mastery-style homework exercises in introductory physics courses: Implementation matters

Brianne Gutmann, Gary Gladding, Morten Lundsgaard, and Timothy Stelzer

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



(Received 30 April 2017; published 18 May 2018)

Encouraged by positive clinical results at the University of Illinois, mastery-style homework was integrated into a large semester-long preparatory physics course via an online homework system that used narrated animated video solutions as correctives. This paper discusses the impact and evolution of the homework in its first two years. The first iteration revealed that students were frustrated and did not engage with the system in an effective way. Intending to reduce that frustration and quell negative behavior, the mastery requirement was relaxed, transfer between versions was reduced, and the addition of a direct discussion with students about the homework were implemented in its second year. The results showed that details of implementation can substantially affect students' behavior; large and statistically significant effects were observed as a reduction in frustration (with self-identified "frustrated" students dropping from 60% in 2014 to 30% in 2015) and improvement in performance (average student mastery rate of 59% to 69%).

DOI: 10.1103/PhysRevPhysEducRes.14.010128

### I. INTRODUCTION

At a large university, an incoming population of students is inevitably going to have a diverse background. This diversity creates an incoming class with a spectrum of familiarity with concepts, abilities in problem solving, and fundamental skills. Despite this, most students in engineering or physics are advised to take the same introductory physics sequence, treated to the same material regardless of the depth of their preparation in high school. While creating course materials, one hopes for an approach that will cater to as many students as possible, while also catching those who may otherwise be at risk to fail.

When educator Benjamin Bloom published his proposal "Learning for Mastery" in 1968 [1], a key assertion that he made was that *any* student can learn, regardless of inherent aptitude and baseline skills, as long as instructional materials and time spent on concepts were adjusted to match each individual's needs [1,2]. He was inspired by Carroll's definition of aptitude as the amount of time required to attain mastery, with the additional caveat that adjusting time allows any person to reach it [3,4]. Rather than giving each student the same treatment, learning could be paced so that students who were comfortable could move on while students who were struggling with specific concepts could spend more time and receive feedback and treatment until they mastered that content.

To implement this idea, Bloom suggested that concepts be broken down into small units; students could be tested on a single unit, then either move on to the next concept if their understanding was sufficient or be given feedback and correctives, if not. Once a student believes they have understood the concept and completed the correctives, they have the opportunity to retest, which can allow them to move onto the next unit, or to receive more practice if the retest shows that the skills have not yet been mastered. This continues until mastery is achieved, and then the student is able to move on to the next unit of material, where they repeat the process. Thus, the testing serves a different purpose than typical testing, which often reinforces the gradient of different students' expected abilities [1].

Although Bloom is generally associated with coining the term *mastery learning*, he was not the only person to advocate this method. Fred Keller had developed his own version of "mastery," the Personalized System of Instruction (PSI), which he also published in 1968 in an article titled "Good-bye, Teacher..." [5], the same year that Bloom's Learning for Mastery (LFM) [1] was published. Keller's system was similar to Bloom's, using small competencies and using testing and retesting, with intervening treatments at different paces for different students. As a psychologist, his work was motivated by operant conditioning, where positive behavior is reinforced by reward and poor behavior is discouraged by a negative consequence [4,6]. By creating repetitive opportunities for formative testing, students received many iterations of positive or negative feedback. Because teachers are physically incapable of giving constant feedback to every individual student, Keller suggested the use of proctors or a computer-based delivery method, which marks a difference between the two approaches. As an

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

educator, Bloom focused on K–12 students and group-based correctives with external deadlines for students to self-pace within, whereas Keller suggested individual, potentially computer-based delivery for high school and college level courses, with absolutely no external pacing. The crux of both these methods, however, was the idea of allowing students to spend as much or as little time necessary to work through explicit objectives via tests and correctives.

The original framework for LFM and PSI, as discussed, was inspired by psychological principles of learning: Carroll's time-on-task and operant conditioning, specifically. Beyond that, mastery learning takes advantage of other cognitive and educational tools. By testing and adaptively retesting, the delivery method exemplifies formative assessment, a tool that Bloom cited in 1968 that is now widely understood to improve learning [7]. Further, the act of testing itself has been shown to improve student performance; the testing effect has been documented in many studies, showing that recalling and encoding information are linked cognitively, and that recalling information often enhances retention and performance more than restudy does [8]. One explanation for this phenomenon is transfer-appropriate processing [9], an aspect of the levels-of-processing cognitive framework which argues that “retention [is] determined by how well the processing requirements of the test matched those used originally to encode information” [10], meaning that the more similar the procedures used to encode and recall information, the better students are able to retain the information. By learning through testing, students are at an advantage to recall information later during testing. The testing effect has been shown most effective when students are given feedback, while Merrienboer *et al.* have demonstrated the benefit of delivering feedback “precisely when learners need it” [11], a method exploited by mastery learning. Considering Vygotsky's zone of proximal development (ZPD), mastery learning should aim to create assessments that are within the zone between what students can do alone, and what they cannot do at all. By giving students tasks that they are able to achieve with help, then providing that help, the students will be learning within their ZPD [12].

Beyond the reasons mastery learning *should* be a useful teaching tool, these principles have been applied in many variations and significant research has looked at the effectiveness of the mastery method in practice. Kulik and Kulik, with various collaborators, have done extensive meta-analyses of studies that identify as Keller's Personalized System of Instruction and Bloom's Learning for Mastery, assessing achievement, retention, attitudes, and types of learning [13–16]. Their most recent, published in 1990 with Bangert-Drowns [13], reevaluated and synthesized one of their earlier works [16] with other meta-analyses done by Guskey and Gates [17] and Slavin [18], evaluating a total of 108 papers. From this analysis, Kulik *et al.* discovered that mastery learning has shown significant improvement over

other traditional systems for students' immediate performance and retention. Additionally, a review by Block and Burns [4] cites mastery's ability to address education debt by helping both low-performing and high-performing students. By making learning objectives concrete, correctives available, and time on task changeable, learning variation is reduced, but not at the expense of high-performing students. Further, nearly all studies that examined student attitudes showed improvement. Block and Burns specifically reported improvement in “interest in and attitudes towards the subject matter learned, self-concept (academic and more general), academic self-confidence, attitudes towards cooperative learning, and attitudes toward instruction” [4]. Additionally, although critics of mastery learning worry that it is only suitable for rigid, algorithmic skills, mastery has been shown to also be effective for teaching critical thinking. Block and Burns assert from their review that mastery can teach “comprehension, application, analysis, synthesis, and evaluation skills” [4]. Mastery style courses have been shown to be well suited for problem solving and higher-order thinking [4,19,20].

Though much of the research using the LFM and PSI has been directed at a variety of courses outside of STEM, there are positive results in mathematics courses, as well as a movement in the 1970s when several universities implemented self-paced introductory physics courses which were modeled after Keller's PSI [21–26]. Although the movement was enthusiastic and college physics seems well suited for mastery, the use of these courses did not become widespread. There have been successful online homework systems with immediate feedback within physics [27–29], but these did not use the mastery delivery system, although they took advantage of some similar learning principles related to online systems with frequent feedback [30–32]. These include the commonly used Mastering Physics system, which is also online with interactive feedback but is not mastery learning as described by Bloom and Keller, despite its similar name [33]. Most of these systems differ from mastery by not providing variation for repeated attempts and correctives. Recently, Heckler and Mikula have implemented an intervention using mastery learning for delivery in the physics domain, but their work focuses on essential skills, specifically vector math as is applied to physics, rather than holistic content. Students in their course spread 2–3 hours of online skills practice over an entire semester, rather than relying solely on mastery homework [34,35]. They describe their work as similar to the online intelligent tutor ALEKS, which also uses mastery delivery for math skills [36].

Intending to utilize mastery learning principles within physics, two clinical trials were performed at the University of Illinois at Urbana-Champaign [37,38]. In both trials, content was delivered through an online homework system, which allowed units to remain locked until a threshold score on the previous unit was achieved. Students worked

until they submitted all answers, then the system graded the set of questions. After grading, multimedia correctives (in the form of conceptually grounded worked solution videos) were available for all questions, regardless of student correctness, and were available to be watched as many times as desired. Once a student was satisfied, they could call up a new version of the same unit (if mastery was not reached) or move on to the next unit (if mastery was shown). Calling up a new version of the same unit removed access to the previous version and solutions, so students had to attempt the new version without prior solutions at hand. However, the last version seen of each completed unit and its solutions remained accessible after students moved on. The students and content were both from the calculus-based introductory electricity and magnetism course, typically taken by second year students; the students participating were volunteers.

The first clinical study suggested that narrated animated solution videos are useful correctives [38]. Besides the benefit of scalability, this type of corrective calls on the advantages of multimedia learning [39] and the worked example effect [40], and is an example of “elaborated feedback” [35], shown to be most effective. Confident in our correctives, the second clinical trial compared the mastery-style delivery system against immediate feedback online homework; although students ended up spending similar amounts of time with content, the post-content assessment showed considerable learning gains (effect size = 0.9 and  $p < 0.005$ ) in favor of students using mastery-style homework over traditional online homework [37].

Encouraged by successful implementation in clinical trials, the next step was to broaden the reach of mastery-style exercises. While mastery-style classes tend to take longer, it has been shown that the proportional time disparity between traditional and mastery delivery methods diminishes as the length of the course increases; the difference is most stark when using mastery for a short time, but is reduced if used for an entire semester [4]. Thus, the application of mastery learning in a semester-long course is ideal, and the takeaway of mastery research, as it stands, is that mastery learning can be very effective for maximizing achievement, retention, positive student attitudes, and higher-order thinking skills for all levels of students.

This paper discusses a specific implementation of mastery-inspired online homework in a large semester-long physics course in its first year, as well as documenting data-driven improvements to the system that were effected in its second year. The original research question was to determine whether positive clinical results would effectively transfer into a classroom, as well as discover which factors needed to be considered for its success. Informed by the first year’s result, our second year’s research question focused more specifically on how to alter the system and guide student interaction to mitigate student frustration and increase their productive engagement.

## II. STUDENT POPULATION

The preparatory course, *Thinking About Physics*, was deliberately chosen as the best candidate to take advantage of mastery learning’s strengths. Offered at the University of Illinois, the course is designed specifically to serve students from diverse educational backgrounds by training them in skills and problem solving necessary to succeed in the introductory physics sequence. Approximately 500 students take the course in the fall semester, recommended because they scored poorly on a local diagnostic test or because they expressed concern about their preparedness. The students in the course are typically first-term freshmen, the majority of whom are engineers required to take the physics sequence.

The course’s structure consists of narrated animations serving as online prelectures and in-person interactive lectures, supplemented by weekly online homework assignments and 20-person discussion classes led by a teaching assistant (TA). In content, *Thinking About Physics* covers the first third of the introductory mechanics course, reviewed more slowly with an emphasis on conceptual grounding and problem solving. Since mastery-style learning has been shown to be effective for both high and low performing students, it is well suited to a course that intends to bring students from different educational backgrounds up to the same target readiness.

## III. FALL 2014: METHODS

Significant research has demonstrated that mastery can be an effective tool, but also advises that details of implementation affect its success. The meta-analysis of mastery studies by Kulik *et al.* [13] cited that studies which had the biggest positive effects shared the following characteristics: locally developed tests (as opposed to national standardized tests), a high mastery requirement (for example, 90% or 100% performance yielding mastery), teacher-paced units (instead of entirely free of time restraints), and effective correctives. We followed these guidelines in implementing mastery-style homework in our course.

Homework content was developed locally by first identifying competencies for each week of instruction. The competencies were informed by emphases placed on past exams, student performance on those exams, and anecdotal experience of student weaknesses within the course. From those competencies, we developed problems to target the learning goals prescribed, creating four similar (but not identical) versions for each level (Each level is a single “unit” of content, as described in the background section.). To clarify, a schematic of example levels and versions within a single week are shown in Fig 1. Typically, there were three or four levels in each assignment, each level consisting of about five questions. The problem difficulty was set so we could define the threshold for mastery at 100%. To prevent frustration from rounding errors or careless mistakes, the questions were given in multiple-choice format, with

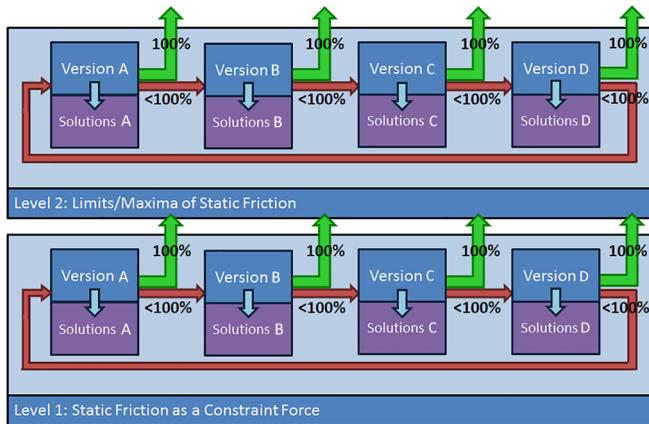


FIG. 1. Online mastery schematic shown of two example levels within week 7: Most weeks had 3–5 levels, with 4 versions (A–D) of problems and solutions each. Students were randomly assigned to version A, B, C, or D for their first try and had access to solutions after grading the version, then were either allowed to move on to the next level (if mastered) or call the next version of the same level (if unmastered). Calling a new version of the same level removed access to solutions and problems of the previous version.

attractive conceptual distractors among the wrong answer choices. Because the homework assignments were weekly, students could work at their own pace, but were not free of time constraints entirely, as recommended by Kulik *et al.* Lastly, our narrated animated worked solution videos were demonstrated to be effective correctives [38]; the videos were provided for each problem, accessible to students after grading that version, regardless of answer correctness. In total, we created about 500 questions and solution videos. These mastery exercises made up nearly all of the regular homework, completed online by all students during the first seven weeks of the course.

#### IV. FALL 2014: RESULTS

Because the mastery-style homework was administered entirely online, data were automatically collected as students interacted with the system. The most striking results from the semester spoke to student attitudes and behavior; it is not surprising that first-term students over an entire semester would engage differently than volunteers for a single clinical trial, but the gravity of that difference was unexpected. Most students began the term with positive attitudes and used the homework as it was designed, but their frustration grew dramatically and useful behaviors deteriorated over the semester.

Given the great effort that was put into creating objectives, content, and solutions, students' strong emotional response was disappointing. A survey given in the first and fifth weeks asked students to rate the homework on a Likert scale from “very encouraging” to “very frustrating” (Fig 2). The data and written responses showed that there was a big shift in how the students felt about the mastery system.

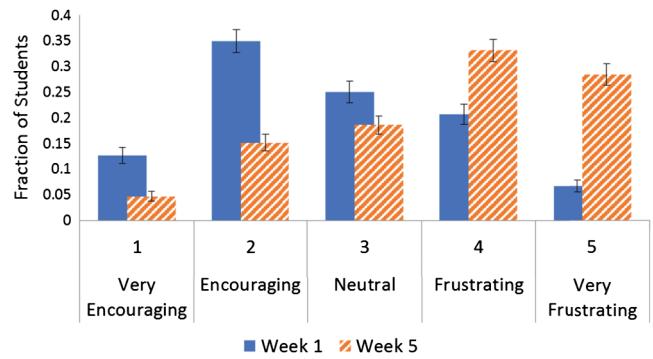


FIG. 2. Normalized frustration histogram for 2014 tested over time: Students responded to the question, “Which statement best reflects your experience working through the mastery homework?”

For example, a student told us on the survey in week 1: “For as many times as I may fail, being able to keep working at new problems until I fully grasp the concept is very encouraging. I feel that I have already learned so much just from the mastery homework.” The same student in week 5 told us, simply, “I don’t like that if you get one problem wrong you have to redo the entire thing.” The percentage of students who identified as frustrated or very frustrated doubled from about 30% to 60% over these weeks. A similar level of frustration was not reported as typical in other online physics homework systems [27–29].

This increased frustration led many students to adopt nonproductive homework behaviors in at least two specific areas that we observed. First, we found that most students spent little time using the provided solution videos. For example, only 40% of the students viewed solutions to at least half of the questions that they answered incorrectly, where we defined “viewing” a solution as watching at least 10% of its total narration time. (Analysis was also done with more strict definitions of viewing, but results did not change much, only decreased the statistics.) This is in contrast to the behavior observed in the clinical trial where over 95% of the students watched all of the relevant solutions. Second, we observed that as the term went on and the difficulty of the levels increased, many students were actually spending less time answering all the questions in a level in a given try. We believe that this behavior arises from the use of one of two nonproductive strategies. Since a given level has four unique versions which cycle, one strategy would be for students to study the solutions to the first version they were given, then choose answers quickly for next three versions to return and apply what they had learned from the first version’s solutions to the repeated version (albeit with different numbers) on their fifth try. The other strategy would be to try to take advantage of the multiple-choice format by simply choosing answers at random to questions that they had no idea how to answer. We can identify both strategies by looking at the average amount of time students spent per try on their second through fourth tries. In particular, we defined the

students engaging in these alternative strategies as those who spent less than a threshold amount of time, defined for each level. This was most often 60 sec, but ranged down to 30 sec for quicker levels. Each particular threshold was determined by examining distributions for students that mastered the level before try five. The fraction of students engaging in these nonproductive strategies was less than 5% for the first couple of weeks, when the material was easier, but increased to peak at 30% for the most difficult material in the final week of the term.

## V. FALL 2014: DISCUSSION

The primary takeaway from the first year of implementing mastery-style homework was that if we want the students to use the materials productively, affect needs to be addressed directly. Because the results from our clinical trial had no such issues, we neglected to consider students' emotional response to the system while dealing with it for an entire semester. Many students were not using the homework as intended, but rather circumventing the learning portions (sincere attempts at the problems and the solutions) in order to get their points as quickly as possible. Though this behavior is not completely unexpected for first-semester college students, the scale and impact of student frustration and desperation was greater than anticipated. As the term progressed, these issues compounded; students became frustrated, used the system in ways that did not help them learn, and then felt more frustrated when they could not demonstrate learning. To address this, we needed to look for ways to improve student attitudes about mastery-style homework to encourage them to use the system in a useful way.

A potential cause of frustration was mismatch among versions of the same level, as far as content and difficulty. Wording or coincidence among multiple-choice answers sometimes made specific questions harder or easier than in other versions, which led to either frustration by passing students before they were prepared for the following level or frustration because the practice and solution videos did not prepare them for an unusually difficult problem. Further, scenarios that we considered equivalent (for example, free body diagrams on a ramp with rotated coordinate axes versus free body diagrams with traditional axes) had vast performance differences. One reason that students cited for their lack of interest in the solution videos was that they felt that each new version was unrelated to the last, and thus had little motivation to review missed questions.

Moving forward, it was clear that affect needed to also be addressed directly. General student attitudes towards the homework were resentful, which may have also been due to students' perceived lack of agency. The language used by students portrayed the situation as something "done to them" as opposed to something "for them," which did not promote students to take ownership of their learning. Rather than a tool, the homework was seen as an obstacle to overcome by any means. Our main objective for the

following year was to reduce frustration and shift student attitudes and behavior.

## VI. FALL 2015: METHODS

The effort to improve student engagement began with a change to the homework's delivery method. To quell students' tendency for skipping intermediate versions or resorting to guessing, students were allowed only four tries per level, meaning they could not cycle back to their original version. Because we restricted their attempts, grading was also adjusted; instead of getting full points for 100% performance and requiring mastery, each student received the best score of their four attempts. By doing this, students were deliberately making a choice to try again for a better score but were not penalized if they tried again and performed less well. Thus, the threshold for moving onto the next level was relaxed to be 100% performance *or* attempting all four versions. As a consequence, those who could not master a level were still able to see later levels and receive some points without resorting to unproductive behavior. Additionally, single weekly homework assignments were split into two smaller biweekly assignments to reduce student fatigue and encourage time management.

To confront student attitudes directly, we included an activity in the first discussion class that asked students to consider and record pros and cons of different types of homework styles, including mastery. Students worked individually and then shared ideas in a class-wide discussion. By allowing students to come up with reasons that the mastery homework could be useful to them and voice potential frustrations, the activity aimed to impress the function of homework as a learning tool, created to help them.

The homework's content itself was modified to create more uniform versions across each specific level. By looking at student performance data on the question, version, and unit level, we identified and adjusted problems that were unintentionally too difficult or too easy. Moreover, new levels were added or old levels were split to target student weaknesses revealed by the data. This served the double purpose of helping students with difficult topics and ensuring that no students could pass through the homework without confronting these weaknesses. Over seven weeks, 27 content levels from Fall 2014 became 32 levels in Fall 2015, and 474 questions were expanded into 620 questions. A list of levels for both 2014 and 2015 is listed in Table I in Appendix A. Two weeks (week 3 and week 7) were left unchanged to enable comparison from the first year to the second. Additionally, feedback was added for the most commonly chosen wrong answers, so that instead of only showing the correct way to work a problem, we could also target students' specific conceptual mistakes.

## VII. FALL 2015: RESULTS

The main objectives for the implementation of mastery-style homework in its second year were to ease frustration

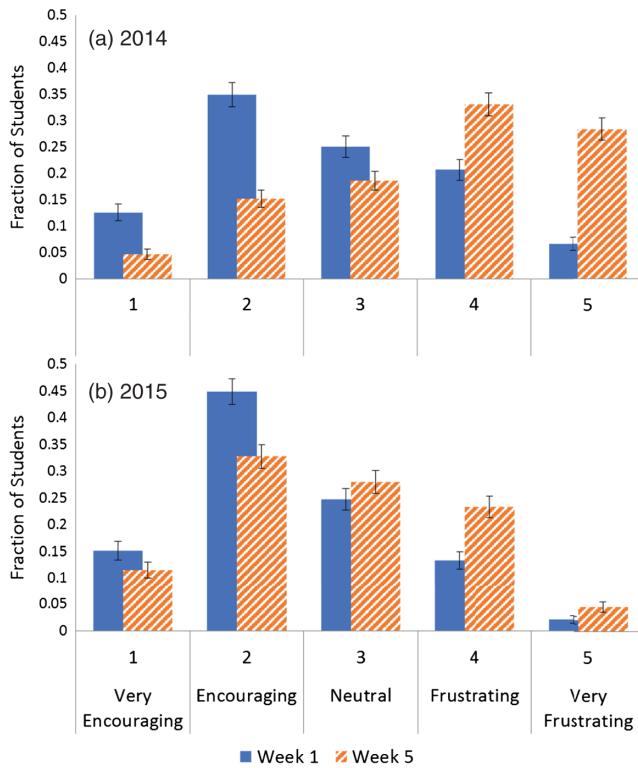


FIG. 3. Normalized frustration histograms from (a) Fall 2014 and (b) Fall 2015: Students responded to the question, “Which statement best reflects your experience working through the mastery homework?”

and improve student engagement with the system. The effect of the changes to the content and delivery dramatically reduced the students’ tendency to become more frustrated as the term went on, as well as improved students’ original perception of mastery-style homework. Figure 3 shows the histogram of responses to the same survey in 2014 and 2015. Though students still found the system more frustrating as time went on, the shift was much less dramatic. By week 5, the portion of students who reported feeling frustrated was only about 30%, halved

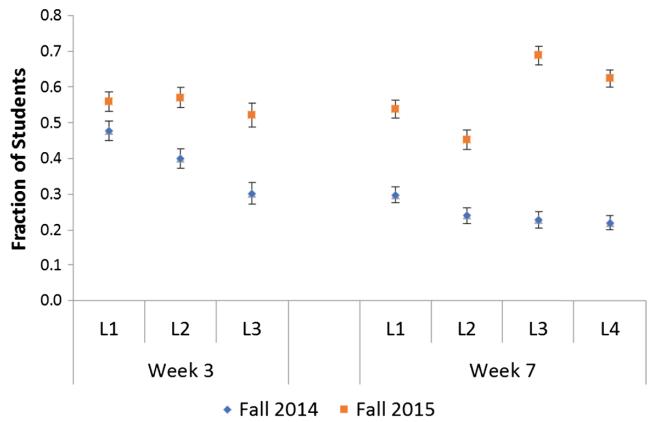


FIG. 4. Solution viewing: The fraction of students who watched over half of the solution videos to their missed problems, separated by levels in two unchanged weeks from 2014 to 2015.

from 60% in 2014. Additionally, students began the semester feeling less frustrated than their classmates did in 2014. The improvement in students’ initial impression of the homework could be attributed to the priming exercise in the first class. Using the average reported scores, the effect size in reducing frustration from the first year to the second year was  $p < 0.001$ .

One expects that from improved attitudes, improved behaviors will follow. Time spent on solution videos significantly increased compared to the previous year; Fig. 4 shows the fraction of students who watched more than 50% of the solutions to missed questions on the unchanged weeks. The threshold for viewing was again 10% of the solutions’ narration time. Although the content was not changed for weeks 3 and 7, students in Fall 2015 were more conscientious about using correctives. The effect is most noticeable in week 7, where the average fraction of students choosing to watch solutions more than doubled. Likewise, using the same thresholds defined previously for students’ average time spent on tries 2–4, less than 5% of students used the unproductive behaviors of guessing or skipping through on week 7, as compared to 30% in Fall 2014.

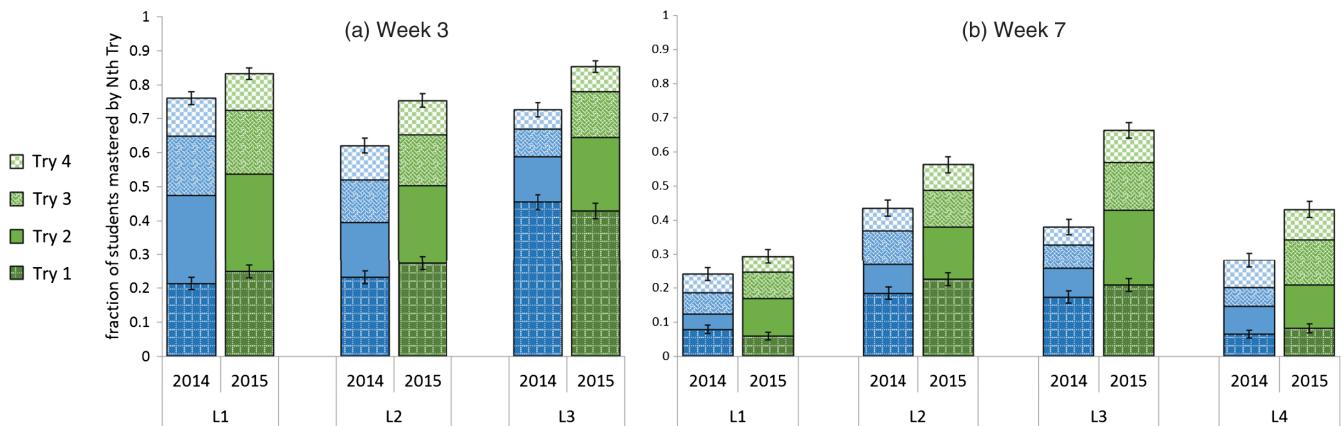


FIG. 5. Student progression through unchanged weeks’ levels in 2014 and 2015, (a) week 3 and (b) week 7: Progression is shown as the fraction of students who had mastered each level by try.

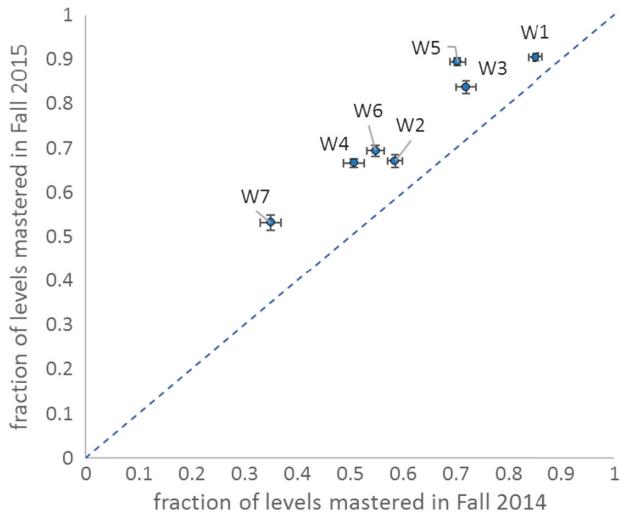


FIG. 6. Average student mastery rates through semester by week: Both axes are the average fraction of levels mastered in each week, in 2014 and 2015. Points above the line indicate improved mastery rates for 2015 over 2014. This plot includes only students who attempted all levels, or all levels except one.

Looking at the number of attempts required for students to master levels (100% performance) also provides information of how students are interacting and learning from the activities. The culmination of the changes made affected student behavior, which is reflected in more serious attempts by students during the second year of implementation. Figure 5 shows the fraction of students who mastered each level by attempt for the two weeks with identical content in 2014 and 2015. Note that the mastery rate on the first attempt is consistent between the two years, suggesting that the groups of students are comparable. For example, the average difference between first attempt scores from 2014 to 2015 in week 3 was  $1.7\% \pm 1.8\%$ , and  $1.8\% \pm 1.1\%$  in week 7. The differences in the groups' baseline understanding were not significant, but students mastered levels much more rapidly in 2015.

Evaluating student performance over the entire semester, the average fraction of levels mastered by students (in the first four tries, before repeated versions in 2014) was systematically higher in 2015 than in 2014, shown as a function of week in Fig. 6. The scatter plot shows the weeks' average fraction of mastering students with Fall 2014 on the horizontal axis and Fall 2015 on the vertical axis; every point is significantly above the line, which indicates improvement of students in 2015 over students in 2014. Integrating over the first seven weeks, this effect is large and statistically significant with effect size of 0.7 and  $p < 0.001$ .

### VIII. CONCLUSIONS

The deliberate changes to the online master-style homework increased students' useful interaction with the system from its first year to its second year. On a small

scale, disparities between versions were smoothed out by a combination of content reorganization and the adjusting of unintentionally problematic questions. These changes, as well as the format change to a finite number of attempts, resulted in students' tendency to engage with solution videos more often, suggesting that the increased homogeneity of the versions made the solutions more useful and the delivery method made each attempt more valuable. The direct affect conversation improved students' initial attitudes towards mastery and the attitudes continued to be more positive, signaling an increase in ownership and appreciation of learning. Students were able to receive partial credit, but worked more sincerely for points even without the threat of receiving none for an unmastered unit.

Unfortunately, it is not possible to isolate the impact of the individual changes. Unlike a clinical study where one has the luxury of prioritizing the research aspect, in this case we needed to prioritize our students. This required us to change all aspects that were likely to improve their learning. However, the data did allow us to infer that each of the components had some impact. The reduction in baseline frustration from one year to the next suggests that our direct affect conversation improved initial student perception of mastery-style homework. A comparison of weeks 3 and 7 (which did not change from 2014 to 2015) showed that students watched more solutions and mastered more often even when content was unchanged. This suggests that students were taking the homework more seriously, rather than the material being simply more manageable. Still, this change in student attitude, behavior, and performance could be a consequence of the delivery style, our affect exercise, or a culmination of content changes in previous weeks. Technically, limiting students to four attempts deviates from pure mastery. However, by reducing the system's strictness, students felt a more compelling sense of agency and performed more as mastery theory prescribes, demonstrated by students' higher rate of mastering levels and using correctives (in our case, video solutions). Despite the relaxation of the mastery requirement, the changes amplified students' positive behaviors that take advantage of mastery learning principles.

When implementing mastery-style homework, the audience and its affective choices are crucial to consider. What was very successful for mature students in our clinical trial was much less successful for novices in a semester-long course. By targeting specific weaknesses, incentivizing positive behavior, and acknowledging student limitations, the system was more appealing to its users and thus, students engaged with it in a more productive way.

### ACKNOWLEDGMENTS

This work was partially supported by the NSF through the IUSE program Grant No. DUE 16-08002.

## APPENDIX A: LIST OF LEVELS IN 2014 AND 2015

TABLE I. List of levels from 2014 and 2015. Levels were given as one assignment in 2014 and broken into two biweekly assignments for 2015.

Topic	Week	2014	2015
Kinematic definitions	Week 1a	L1: $v$ - $t$ graphs L2: $x$ - $t$ graphs L3: Multiple-body graphs	L0: Warmup L1: $v$ - $t$ graphs L2: $x$ - $t$ graphs
	Week 1b	N/A	L3: Math supplement: Coordinate geometry L4: Multiple-body graphs
Constant acceleration and 1D relative motion	Week 2a	L1: Relative motion in 1D L2: 1D Constant acceleration 1 L3: 1D Constant acceleration 2 L4: 1D Constant acceleration 3 L5: 1D Constant acceleration 4	L1: 1D relative motion L2: Constant acceleration L3: Free fall
	Week 2b	N/A	L4: Math supplement: Coupled equations L5: Constant acceleration with symbols
Vectors and 2D relative motion	Week 3a	L1: Vectors L2: 2D Relative motion 1 L3: 2D Relative motion 2	L1: Math supplement: Vectors
	Week 3b	N/A	L2: 2D Relative motion 1 L3: 2D Relative motion 2
Projectile motion	Week 4a	L1: Projectile motion 1 L2: Projectile motion 2 L3: Projectile motion 3	L1: Projectile motion 1 L2: Projectile motion 2
	Week 4b	N/A	L3: Math supplement: Algebraic relationships L4: Projectile motion 3
Newton's second law	Week 5a	L1: Drawing FBDs L2: Equations from FBDs 1 L3: Equations from FBDs 2 L4: Full problem	L1: Drawing FBDs L2: Equations from FBDs L3: Math supplement: Rotated coordinate systems
	Week 5b	N/A	L4: Weight vector in a rotated coordinate system L5: Equations from FBDs (Ramps) L6: Full problem
Newton's first and third laws	Week 6a	L1: Identifying Newton's third law pairs L2: Newton's third law (Static) L3: Newton's third law (Accelerating) L4: Introduction to friction (as a constraint force)	L1: Math supplement: systems of equations L2: Newton's second law for two objects L3: Identifying Newton's third law pairs
	Week 6b	N/A	L4: Newton's third law (Static) L5: Newton's third law (Accelerating) L6: Introduction to Friction (as a constraint force)
Friction and uniform circular motion	Week 7a	L1: Static friction L2: Kinetic friction L3: Static or kinetic friction? L4: Uniform circular motion	L1: Static friction L2: Kinetic friction
	Week 7b	N/A	L3: Static or kinetic friction? L4: Uniform circular motion

## APPENDIX B: EXAMPLE VERSIONS OF A SINGLE LEVEL

The following are multiple versions (A–D) of week 7, Level 1: Static friction, which was the same in 2014 and 2015. The competency for the level is “To be able to determine direction and magnitude of static friction force between objects when they are or are not on the verge of moving relative to one another.” Numeric answer choices are shown on version A but not the following version examples.

### 1. Version A

Two blocks (of masses  $M_1 = 5 \text{ kg}$  and  $M_2 = 8 \text{ kg}$ ) are stacked on top of each other and are resting on a horizontal, frictionless floor. A horizontal force  $\mathbf{F} = 25 \text{ N}$  is applied, as shown, and the blocks move together due to the friction between  $M_1$  and  $M_2$ . The coefficient of static friction between the blocks is  $\mu_s = 0.3$ . A diagram of this scenario is provided in Fig. 7.

1) Which of these is a correct free-body diagram for  $M_2$ ? Answer choices are given in Fig. 8.

2) What is the magnitude of the friction force  $\mathbf{f}_{1 \text{ on } 2}$ ?

- $\mathbf{f}_{1 \text{ on } 2} = 25 \text{ N}$
- $\mathbf{f}_{1 \text{ on } 2} = 23.5 \text{ N}$
- $\mathbf{f}_{1 \text{ on } 2} = 14.7 \text{ N}$
- $\mathbf{f}_{1 \text{ on } 2} = 38.3 \text{ N}$

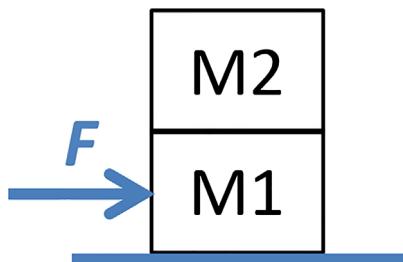


FIG. 7. Diagram provided in the prompt for version A of week 7, level 1 problems.

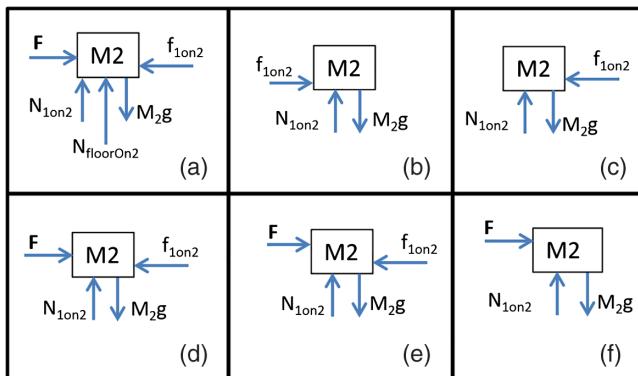


FIG. 8. Answer choices given for question 1 on version A of week 7, level 1 problems.

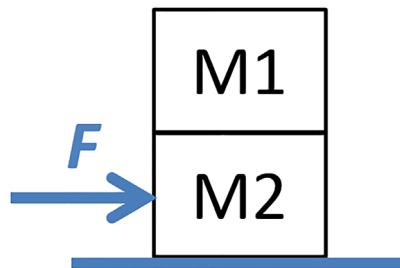


FIG. 9. Diagram provided for question 5 on version A of week 7, level 1 problems.

- $\mathbf{f}_{1 \text{ on } 2} = 9.62 \text{ N}$
- $\mathbf{f}_{1 \text{ on } 2} = 15.4 \text{ N}$
- 3) What is  $\mathbf{f}_{\text{max}}$ , the maximum friction force between the blocks before they start to slip?

  - $\mathbf{F}_{\text{max}} = 38.3 \text{ N}$
  - $\mathbf{F}_{\text{max}} = 14.7 \text{ N}$
  - $\mathbf{F}_{\text{max}} = 9.62 \text{ N}$
  - $\mathbf{F}_{\text{max}} = 25 \text{ N}$
  - $\mathbf{F}_{\text{max}} = 23.5 \text{ N}$
  - $\mathbf{f}_{\text{max}} = 15.4 \text{ N}$

- 4) What is  $\mathbf{F}_{\text{max}}$ , the maximum force  $\mathbf{F}$  that can be applied, so that the blocks will still move together (the top block won't slip)?

  - $\mathbf{F}_{\text{max}} = 38.3 \text{ N}$
  - $\mathbf{F}_{\text{max}} = 15.4 \text{ N}$
  - $\mathbf{F}_{\text{max}} = 25 \text{ N}$
  - $\mathbf{F}_{\text{max}} = 9.62 \text{ N}$
  - $\mathbf{F}_{\text{max}} = 14.7 \text{ N}$
  - $\mathbf{F}_{\text{max}} = 23.5 \text{ N}$

- 5) If we exchange the positions of  $M_1$  and  $M_2$ , how does the magnitude of  $\mathbf{F}_{\text{max}}$  change? A diagram of this new scenario is provided in Fig. 9.

  - $|\mathbf{F}_{\text{max}}|$  decreases.
  - $|\mathbf{F}_{\text{max}}|$  stays the same.
  - $|\mathbf{F}_{\text{max}}|$  increases.

### 2. Version B

A truck of mass  $\mathbf{M} = 15 \text{ kg}$  is accelerating with  $\mathbf{a} = 1 \text{ m/s}^2$  up a hill that is inclined at  $\theta = 17^\circ$  with respect to the horizontal. A crate of mass  $\mathbf{m} = 6 \text{ kg}$  is at rest with respect to the bed of the truck (it is not slipping). The coefficient of

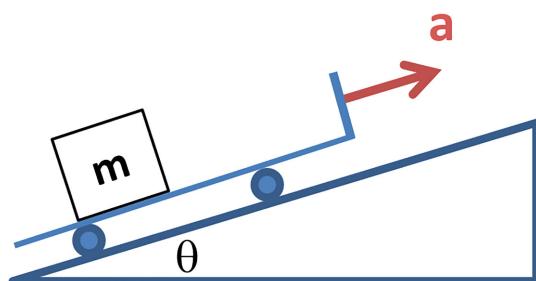


FIG. 10. Diagram provided in the prompt for version B of week 7, level 1 problems.

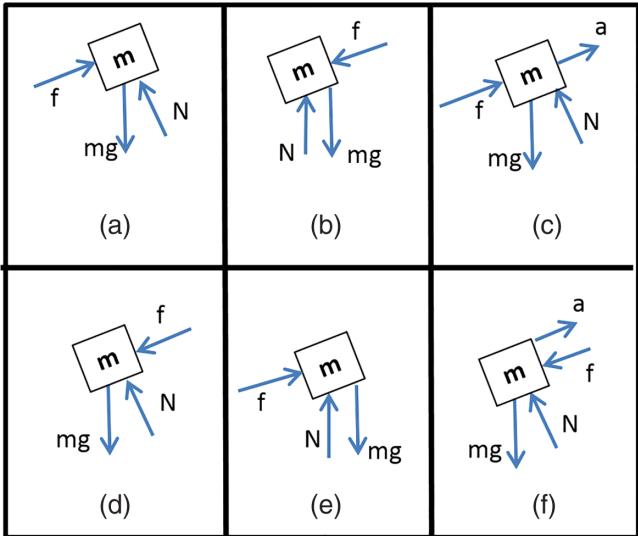


FIG. 11. Answer choices given for question 1 on version B of week 7, level 1 problems.

static friction between the crate and the truck is  $\mu_s = 0.45$ . A diagram of this scenario is provided in Fig. 10.

- 1) Which of these is a correct free-body diagram for the crate? Answer choices are given in Fig. 11.
- 2) What is the magnitude of the friction force on the crate?
- 3) What is  $f_{\max}$ , the maximum friction force between the crate and truck before the crate starts to slip?
- 4) What is  $\mathbf{a}_{\max}$ , the maximum acceleration the truck can have so that the crate doesn't slip?
- 5) How does the magnitude of  $\mathbf{a}_{\max}$  when the truck is accelerating downhill compare to  $\mathbf{a}_{\max}$  when the truck is accelerating uphill?

- a.  $|\mathbf{a}_{\max, \text{downhill}}| < |\mathbf{a}_{\max, \text{uphill}}|$
- b.  $|\mathbf{a}_{\max, \text{downhill}}| > |\mathbf{a}_{\max, \text{uphill}}|$
- c.  $|\mathbf{a}_{\max, \text{downhill}}| = |\mathbf{a}_{\max, \text{uphill}}|$

### 3. Version C

Two blocks of mass  $M_1 = 5 \text{ kg}$  and  $M_2 = 8 \text{ kg}$  are stacked on top of each other and resting on a horizontal, frictionless floor. A force  $\mathbf{F} = 36 \text{ N}$  is applied at an angle  $\theta = 35^\circ$  with respect to the horizontal axis, as shown, and the blocks start moving together. The coefficient of static friction between the  $M_1$  and  $M_2$  is  $\mu_s = 0.4$ . A diagram of this scenario is provided in Fig. 12.

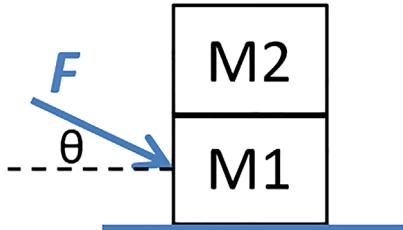


FIG. 12. Diagram provided in the prompt for version C of week 7, level 1 problems.

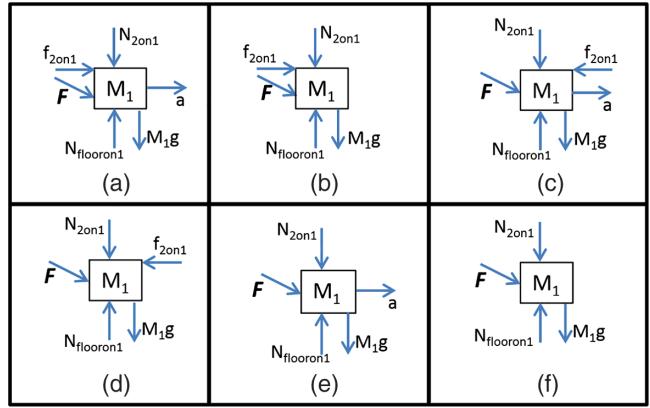


FIG. 13. Answer choices given for question 1 on version C of week 7, level 1 problems.

- 1) Which of these is a correct free-body diagram for  $M_1$ ? Answer choices are given in Fig. 13.
- 2) What is the magnitude of the friction force,  $f_{1 \text{ on } 2}$ , the friction from  $M_1$  on  $M_2$ ?
- 3) What is the maximum magnitude that the friction force,  $f_{1 \text{ on } 2}$  (the friction from  $M_1$  on  $M_2$ ), can have such that  $M_2$  does not slip on  $M_1$ ?
- 4) What is  $\mathbf{F}_{\min}$ , the magnitude of the minimum applied force needed to make  $M_2$  slip on  $M_1$ ?
- 5) If the applied force were applied to  $M_2$  instead of  $M_1$  (with  $M_2$  still on top of  $M_1$ ), how would the magnitude of the friction force between the blocks change, if at all?
  - a.  $|\mathbf{f}_{1 \text{ on } 2}|$  remains the same
  - b.  $|\mathbf{f}_{1 \text{ on } 2}|$  decreases
  - c.  $|\mathbf{f}_{1 \text{ on } 2}|$  increases

### 4. Version D

Two blocks,  $M_1$  and  $M_2$  are stacked on top of each other and resting on a frictionless ramp as shown in the figure. A force  $\mathbf{F}$  is applied to  $M_1$  parallel to the ramp. The blocks move together as they accelerate up the ramp.

The blocks have masses  $M_1 = 10 \text{ kg}$ ,  $M_2 = 4 \text{ kg}$ , and the applied force is  $\mathbf{F} = 65 \text{ N}$ . The ramp's angle of elevation is  $\theta = 24^\circ$ , and the coefficient of static friction between the two blocks is  $\mu_s = 0.6$ . A diagram of this scenario is provided in Fig. 14.

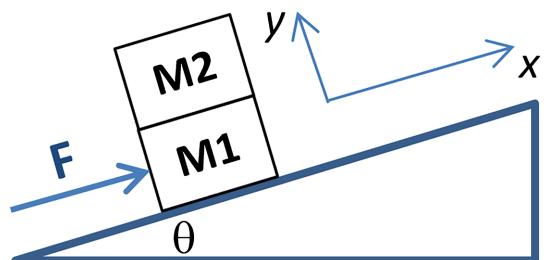


FIG. 14. Diagram provided in the prompt for version D of week 7, level 1 problems.

1) Which of these is a correct free-body diagram for  $M_1$ ? Answer choices are given in Fig. 15.

2) Using the  $x$ - $y$  coordinates set up in the figure, what is the  $x$  component of acceleration of  $M_1$ ?

3) What is the  $x$  component of the friction force  $f_{1\text{on}2}$ ?

4) What is  $F_{\max}$ , the maximum force  $F$  that can be applied so that the blocks still move together?

5) Compare the net force on  $M_1$  to the net force on  $M_2$ , when the applied force is still only the original  $F$ .

- a.  $|\mathbf{F}_{\text{net}}(M_1)| = |\mathbf{F}_{\text{net}}(M_2)|$
- b.  $|\mathbf{F}_{\text{net}}(M_1)| < |\mathbf{F}_{\text{net}}(M_2)|$
- c.  $|\mathbf{F}_{\text{net}}(M_1)| > |\mathbf{F}_{\text{net}}(M_2)|$

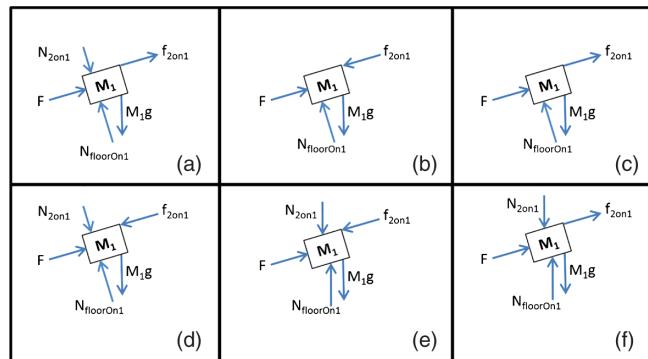


FIG. 15. Answer choices given for question 1 on version D of week 7, level 1 problems.

[1] B. Bloom, Learning for mastery, *Eval. Comment* **1**, 1 (1968).

[2] T. R. Guskey, Closing achievement gaps: Revisiting Benjamin S. Bloom's "Learning for Mastery", *J. Adv. Acad.* **19**, 8 (2007).

[3] J. B. Carroll, A model of school learning, *Teachers Coll. Rec.* **64**, 723 (1963).

[4] J. H. Block and R. B. Burns, Mastery learning, *Rev. Res. Educ.* **4**, 3 (1976).

[5] F. S. Keller, Good-bye teacher..., *J. Appl. Behav. Anal.* **1**, 79 (1968).

[6] B. F. Skinner, Teaching machines: From the experimental study of learning come devices which arrange optimal conditions for self-instruction, *Science* **128**, 969 (1958).

[7] P. Black and D. Wiliam, Assessment and classroom learning, *Assess. Educ. Princ. Pol. Pract.* **5**, 7 (1998).

[8] F. M. Zaromb and H. L. Roediger III, The testing effect in free recall is associated with enhanced organizational processes, *Mem. Cogn.* **38**, 995 (2010).

[9] C. D. Morris, J. D. Bransford, and J. J. Franks, Levels of processing versus transfer appropriate processing, *J. Verbal Learn. Verbal Behav.* **16**, 519 (1977).

[10] H. L. Roediger, III, D. A. Gallo, and L. Geraci, Processing approaches to cognition: The impetus from the levels-of-processing framework, *Memory* **10**, 319 (2002).

[11] J. J. G. van Merriënboer, P. A. Kirschner, and L. Kester, Taking the load off a learner's mind: Instructional design for complex learning, *Educ. Psychol.* **38**, 5 (2003).

[12] L. S. Vygotsky, *Mind and Society: The Development of Higher Mental Processes* (Harvard University Press, Cambridge, MA, 1978).

[13] C.-L. C. Kulik, J. A. Kulik, and R. L. Bangert-Drowns, Effectiveness of mastery learning programs: A meta-analysis, *Rev. Educ. Res.* **60**, 265 (1990).

[14] J. A. Kulik, K. Carmichael, and C.-L. Kulik, The Keller plan in science teaching: An individually paced, student-tutored, and mastery-oriented instructional method is evaluated, *Science* **183**, 379 (1974).

[15] J. A. Kulik and C.-L. C. Kulik, Effectiveness of the personalized system of instruction, *Eng. Educ.* **66**, 228 (1975).

[16] J. A. Kulik, C. C. Kulik, and P. A. Cohen, A meta-analysis of outcome studies of Keller's personalized system of instruction, *Am. Psychol.* **34**, 307 (1979).

[17] T. R. Guskey and S. L. Gates, A synthesis of research on group-based mastery learning programs, *Proceedings of the Annual Meeting of American Education Research Association in Chicago, IL* (1985) (unpublished), <https://eric.ed.gov/?id=ED262088>.

[18] R. E. Slavin, Mastery learning reconsidered, *Rev. Educ. Res.* **57**, 175 (1987).

[19] Z. Mevarech and S. Werner, Are mastery learning strategies beneficial for developing problem solving skills?, *Higher Educ.* **14**, 425 (1985).

[20] K. Spencer, Media and technology in education: Raising academic standards (1996).

[21] O. T. Anderson and R. A. Artman, A self-paced, independent study, introductory physics sequence—description and evaluation, *Am. J. Phys.* **40**, 1737 (1972).

[22] M. A. Philippas and R. W. Sommerfeldt, Keller vs lecture method in general physics instruction, *Am. J. Phys.* **40**, 1300 (1972).

[23] B. A. Green, Jr, Physics teaching by the Keller plan at MIT, *Am. J. Phys.* **39**, 764 (1971).

[24] G. C. Brown, K. I. Greisen, and A. J. Sievers, Self-paced introductory physics course—an eight-year progress report, *Am. J. Phys.* **45**, 1082 (1977).

[25] G. D. Putt, Testing the mastery concept of self-paced learning in physics, *Am. J. Phys.* **45**, 472 (1977).

[26] J. L. Safko, Structure for a large enrollment, self-paced, mastery oriented astronomy course, *Am. J. Phys.* **44**, 658 (1976).

[27] 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).

[28] E.-S. Morote and D. Pritchard, Technology closes the gap between students' individual skills and background differences, in *Proc. Soc. Inf. Technol. Teach. Educ. Int. Conf. 2004* edited by R. Ferdig, C. Crawford, R. Carlsen, N. Davis, J. Price, R. Weber, and D. A. Willis (Association for the Advancement of Computing in Education (AACE), Atlanta, GA, USA, 2004), pp. 826–831.

[29] 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).

[30] W. Wager and R. W. Kulhavy, Feedback in programmed instruction: historical context and implications for practice, in *Interactive Instruction and Feedback*, edited by J. V. Dempsey and G. Ales (Educational Technology Publications, Englewood Cliffs, NJ, 1993), pp. 3–20.

[31] J. Hattie and H. Timperley, The power of feedback, *Rev. Educ. Res.* **77**, 81 (2007).

[32] F. M. Van der Kleij, R. C. W. Feskens, and T. J. H. M. Eggen, Effects of feedback in a computer-based learning environment on students' learning outcomes, *Rev. Educ. Res.* **85**, 475 (2015).

[33] Mastering PhysicsPearson, <http://www.pearsonmylabandmastering.com/masteringphysics/> (2017).

[34] 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).

[35] A. F. Heckler and B. D. Mikula, Factors affecting learning of vector math from computer-based practice: Feedback complexity and prior knowledge, *Phys. Rev. Phys. Educ. Res.* **12**, 010134 (2016).

[36] W. Canfield, ALEKS: A web-based intelligent tutoring system, *Math. Comput. Educ.* **35**, 152 (2001).

[37] G. Gladding, B. Gutmann, N. Schroeder, and T. Stelzer, Clinical study of student learning using mastery style versus immediate feedback online activities, *Phys. Rev. ST Phys. Educ. Res.* **11**, 010114 (2015).

[38] N. Schroeder, G. Gladding, B. Gutmann, and T. Stelzer, Narrated animated solution videos in a mastery setting, *Phys. Rev. ST Phys. Educ. Res.* **11**, 010103 (2015).

[39] R. E. Mayer, *The Cambridge Handbook of Multimedia Learning* (Cambridge University Press, Cambridge, England, 2005).

[40] J. Sweller, The worked example effect and human cognition, *Learn. Instr.* **16**, 165 (2006).