



Faculty & Staff Scholarship

2016

Behavioral Self-Regulation in a Physics Class

John Stewart

Seth DeVore

Gay Stewart

Follow this and additional works at: https://researchrepository.wvu.edu/faculty_publications

Behavioral self-regulation in a physics classJohn Stewart,^{*} Seth DeVore,[†] and Gay Stewart[‡]*Department of Physics and Astronomy, West Virginia University, Morgantown, West Virginia 26506, USA*

Lynnette Michaluk

*Center for Excellence in STEM Education, West Virginia University,
Morgantown, West Virginia 26506, USA*

(Received 10 December 2015; published 1 April 2016)

This study examined the regulation of out-of-class time invested in the academic activities associated with a physics class for 20 consecutive semesters. The academic activities of 1676 students were included in the study. Students reported investing a semester average of 6.5 ± 2.9 h out of class per week. During weeks not containing an examination, a total of 4.3 ± 2.1 h was reported which was divided between 2.5 ± 1.2 h working homework and 1.8 ± 1.4 h reading. Students reported spending 7.6 ± 4.8 h preparing for each in-semester examination. Students showed a significant correlation between the change in time invested in examination preparation ($r = -0.12$, $p < 0.0001$) and their score on the previous examination. The correlation increased as the data were averaged over semester ($r = -0.70$, $p = 0.0006$) and academic year ($r = -0.82$, $p = 0.0039$). While significant, the overall correlation indicates a small effect size and implies that an increase of 1 standard deviation of test score (18%) was related to a decrease of 0.12 standard deviations or 0.9 h of study time. Students also modified their time invested in reading as the length of the textbook changed; however, this modification was not proportional to the size of the change in textbook length. Very little regulation of the time invested in homework was detected either in response to test grades or in response to changes in the length of homework assignments. Patterns of regulation were different for higher performing students than for lower performing students with students receiving a course grade of “C” or “D” demonstrating little change in examination preparation time in response to lower examination grades. This study suggests that homework preparation time is a fixed variable while examination preparation time and reading time are weakly mutable variables.

DOI: 10.1103/PhysRevPhysEducRes.12.010125

I. INTRODUCTION

Total time on task (TOT) is one of the seven principles for good practice in undergraduate education identified by Chickering and Gamson [1]. In 2012 the National Survey of Student Engagement found that of the 108 015 first-year students responding to a question asking for the time spent per week preparing for all classes 36% reported spending 10 h or less per week while only 23% reported spending more than 20 h. These reported time investments changed little for the 150 524 seniors in the study, 37% of whom reported working 10 h or less and 26% who reported working more than 20 h. The responses were also very similar across types of institutions from research universities with a very high research ranking to four-year liberal arts colleges [2]. Out-of-class TOT is also collected by

major international studies of K–12 education such as the Program for International Student Assessment (PISA) and Trends in International Mathematics and Science Study (TIMSS), making TOT one of the most broadly studied educational constructs.

To build understanding in a science class, students engage in a variety of activities both inside and outside the classroom. The amount of time invested in academic activities outside of the classroom and the kind of activities performed accounts for 21% to 36% of the variance in the test averages and for 19% to 37% of the variance in the normalized gains [3] on the Conceptual Survey of Electricity and Magnetism (CSEM) [4]. The amount of variance explained by TOT is comparable to the variance in class performance which is explained by logical reasoning ability (19%[5]–24%[6]), mathematical reasoning ability (12%[5]–26%[7]), or physics pretest score (1%[8]–30%[7]).

In this paper, the total out-of-class TOT invested by students in an introductory physics class will be explored. The total time spent in working homework, reading, and preparing for examinations will be presented. Changes in time investment will be compared for semesters with different assignment lengths. Changes in student time

^{*}jcstewart1@mail.wvu.edu[†]stdevor@mail.wvu.edu[‡]gbstewart@mail.wvu.edu

investment within the semester will also be examined and correlated with examination scores.

Extensive research into the effect of TOT has been performed in many fields, but little work specific to physics or STEM (science, technology, engineering, and mathematics) classes at the college level has been reported [9]; however, a few studies in physics and engineering have been conducted. Di Stefano looked at student behavior in a number of reformed physics classes and found that students spent an average of 6 h per week on a diverse set of out-of-class activities [10]. This time investment is similar to the time spent working online physics homework reported by Kortemeyer [11]. Welch and Bridghan found a negative correlation between the amount of time spent covering a unit of instruction in a physics class and achievement; thus, students who spent more time covering a unit earned lower grades than those who spent less time on the unit [12]. Conversely, Springer *et al.* found effect sizes ranging from 0.52 to 0.63 for the relationship of the time spent in small group activities in a STEM class and achievement in that class [13]. Studies of the overall time commitment of engineering students at a number of institutions show that these students invest substantially less time in their studies than is expected by their institutions [14,15].

Time has also been employed as a variable in a number of more detailed studies. The time spent by experts classifying item difficulty was used to explore expert or novice perceptions of physics problem difficulty [16]. The time required to detect differences in physics problems has also been used to study expert or novice differences [17]. Time differences in the response to in-class personal response system conceptual physics questions have been used to understand student reasoning [18].

Outside of STEM education, TOT measurements have been used to understand educational systems, both traditional and, more recently, online systems. TOT is also an important control variable when investigating the effectiveness of new educational offerings [19,20] because an educational reform that modifies the students' TOT may be more effective simply because students spend more time on a topic. As an example of the need to control for TOT, a metaanalysis by the U.S. Department of Education found that online learning experiences that required less than or equal the amount of TOT as face-to-face instruction had an effect size of 0.19 while those that required more TOT had an effect size of 0.45 [21]. Within physics, the time to answer online physics questions has been used to detect cheating behavior [22]. The time spent on an online activity has also been used as a control to compare different online patterns of problem presentation [23]. The time spent accessing online resources in an online physics homework system was investigated and gender differences in resource access patterns detected [11].

Beyond the time invested in a class, the effect of performing specific class-related activities, such as working

homework, has been investigated. Goldstein provides a review of the effectiveness of homework in precollege classrooms [24]. In a metaanalysis, Paschal *et al.* found an effect size of 0.36 for the assigning and grading of homework on overall performance [25]. A metaanalysis synthesizing 16 years of research on the effect of assigning and the resulting time invested in homework found that the assignment of homework had a positive effect, but that the time invested in homework showed weaker effects [26]. A positive relation between time spent working homework and achievement was found for German students in analyses of the PISA and TIMSS international studies, but a more careful analysis controlling for differences between schools found that at the student level the amount of time spent working homework was negatively correlated with academic achievement [27]. Homework time was also not significantly correlated with achievement correcting for school characteristics in a multischool Dutch study [28]. The amount of overall study time was also found to be weakly related to overall academic performance measured by grade point average (GPA) [29].

Out-of-class TOT plays a central role in understanding learning. It is recognized by practitioners in scientific disciplines as a necessary prerequisite to achieving mastery of the challenging material presented in physics classes. TOT also plays the role of an effect of many of the most studied aspects of educational theory: self-regulation, metacognition, self-efficacy, motivation, resilience, effort regulation, and time and resource management [30–37]. Out-of-class TOT is closely related to self-regulation and its subfacets of effort regulation and time and resource management. Self-regulation may be time neutral where a student changes the way he or she studies but not the time spent; however, a common self-regulation strategy when not performing well in a task is to increase the time committed to the task. Further, clinically, time neutral self-regulation strategies are limited by TOT, because no amount of change of how one studies can overcome the effect of not investing enough time in one's studies. Studies of self-regulation are often gathered as self-reports on the intention to change behavior; TOT provides a measure of whether that intention was put into practice.

Self-regulation strategies have been found to be significantly positively correlated with course performance measures [33,38] and are assumed to be crucial to academic achievement [39–41]. As it pertains to learning, self-regulation refers to the learner's self-directed processes and self-beliefs to transform their mental abilities into academic performance skills. Self-regulation is seen as a set of proactive processes such as setting goals, selecting and deploying strategies, and self-monitoring one's effectiveness students use to acquire academic skills. One potential outcome of these processes is the modification of TOT. These qualities arise from advantageous learner motivational feelings, beliefs, and metacognitive strategies and are

characteristics of good learners [42]. Students who self-regulate tend to perform better in courses than those who do not [43,44].

Self-regulation is particularly important because it has been shown to be a mutable student variable [45]; students can be taught self-regulation strategies [46]. For example, Shen *et al.* [43] found that online vocational students who received instruction in self-regulated learning strategies including time management, goal setting, and self-evaluation performed better in their online learning course than students that didn't receive training.

Self-regulation has many components; the two most closely affecting TOT are effort regulation and time management. Effort regulation has been investigated as a variable affecting academic performance and influencing the understanding of performance. It has been shown to have substantial explanatory power for student class outcomes [47,48]. It has also been shown to be a mediating variable for the effects of other educational variables such as personality traits [49] and self-efficacy [47] on academic performance. Time management has also been investigated as a self-regulatory strategy and shown to be a significant predictor of student success [50,51] and to be important in the transition to college for some underrepresented populations [52]. Time management was examined as an outcome variable resulting from self-regulatory behavior [53] and experiments teaching students time management skills have shown positive results [54]. A meta-analysis of 190 studies [55] demonstrated that effort regulation was one of the factors most strongly correlated with academic success ($r = 0.32$) and that time management was also relatively strongly correlated (compared to other variables) ($r = 0.22$). In this study, only grade goal and performance self-efficacy were more strongly correlated with academic performance. Effort regulation is also the natural outgrowth of other self-regulatory processes which have been shown to correlate with increased academic performance [56]. For a review of topics in self-regulation, see the collection edited by Zimmerman and Schunk [57] and previous editions in this series.

Studies of self-regulation, effort regulation, and resource or time management have demonstrated a relationship between self-regulatory strategies and academic performance; few of these studies have directly investigated the self-regulatory behavior of science students who often have complex and demanding required assignments and access to a diverse set of self-regulation strategies. In their review of why reform physics educational methods are not adopted with greater frequency, Fraser *et al.* identify a lack of studies of out-of-class time in physics classes; this study begins to explore this important research domain within physics [9]. Further, most studies rely on self-reported student intentions to self-regulate, or student impressions of self-regulation, but do not measure how (and if) these intentions are enacted and to what extent they are enacted.

This study is unique in that time commitment was measured at two distinct points in the same class; this within-subjects design will allow the determination of the extent to which effort regulation results in increased time commitment in response to the feedback provided by examination scores.

This study seeks to answer the following questions: *Do students regulate the amount of time invested in a science class in response to changes in that class or to their performance in the class? Does the regulation of time investment change for students achieving different class outcomes?*

II. THEORETICAL FRAMEWORK

Changes in TOT and the self-regulatory behavior implied by those changes will be viewed through the lens of Zimmerman's theory of self-regulation [58] which was further refined by Pintrich [35]; both frameworks were strongly informed by Bandura's social cognitive theory [59]. This framework posits four stages of self-regulatory learning: planning and goal setting, self-monitoring, effort control and task performance, and reflection. We acknowledge that there are many other facets that warrant study such as the quality of effort or the degree of metacognitive monitoring. Self-regulation in response to internal or external feedback may involve time-neutral changes in learning strategies (as we will see in the regulation of homework time), but may also require directing additional effort at the assigned task, modifying TOT; thus, this study measures one potential external outcome of the self-regulation of learning.

III. RESEARCH METHODS

A. Context for research

The course studied was the second-semester, calculus-based, introductory electricity and magnetism course at a large midwestern land-grant university serving approximately 25 000 students. The classroom component of the course consisted of two 50-min traditional lectures and two 2 h laboratory sessions each week. Homework was collected at the beginning of most lecture sessions and a lecture quiz was given at the end of most sessions in order to manage attendance. The laboratory sessions were a mix of small-group problem-solving activities, hands-on inquiry-based investigations, teaching assistant (TA)-led problem solving, TA-led interactive demonstrations, and traditional experiments. Students received credit for completing each laboratory activity and took a quiz during the laboratory session to test how much of the previous homework assignment was retained. The textbook was written specifically for the course and could be modified by the lead instructor, providing a strong link between the reading assignments and the work in the laboratory component. The lecture and laboratory components were

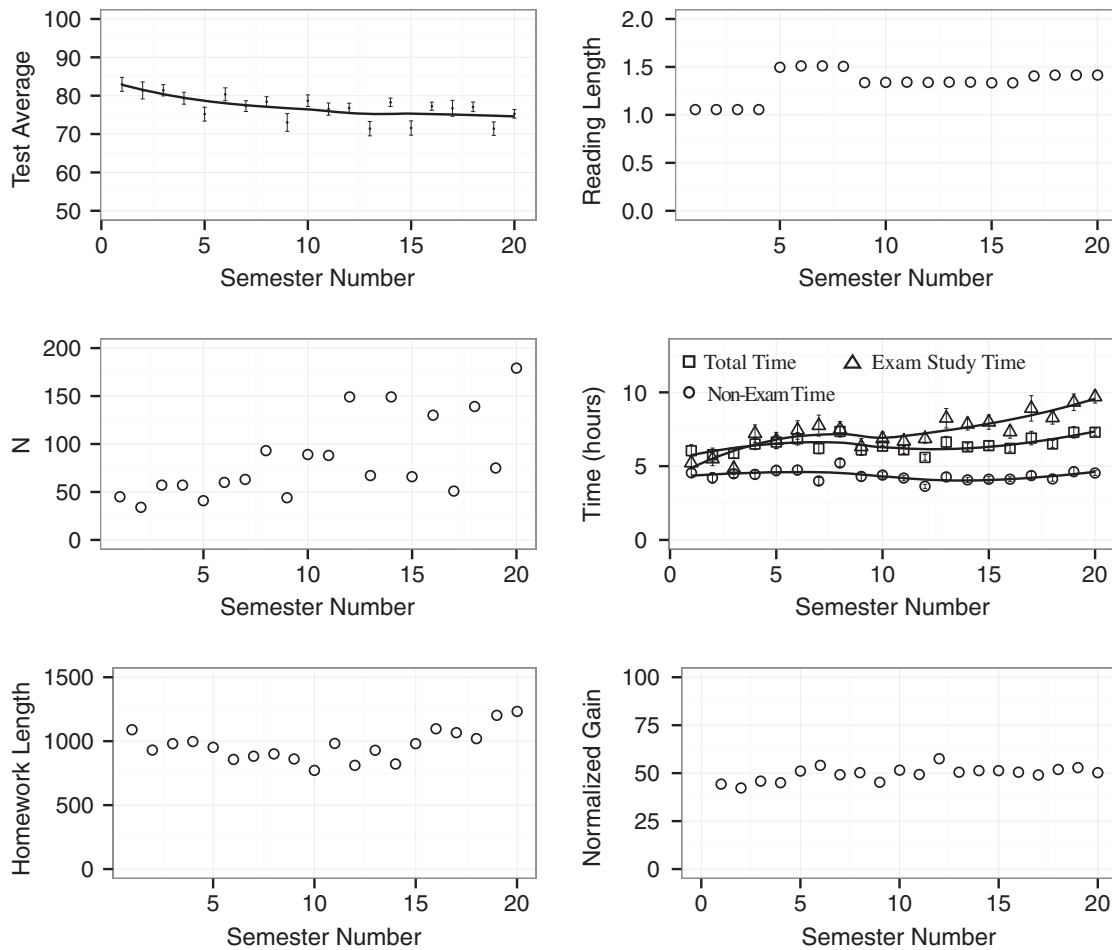


FIG. 1. Variation in class parameters. The average on the first two tests, the number of students included in the study (N), the total length of the homework measured by the number of independent steps in the solutions, the length of the reading measured in the number of characters in millions of characters in the LaTeX file used to generate the course textbook, the normalized gain on the CSEM, and the overall averages of the time variables. All quantities are plotted against the semester number. Semesters are numbered beginning in Fall 2002.

carefully timed so that a lecture often discussed the upcoming laboratory and made use of experiences from previous laboratory sessions.

Learning was monitored with four in-semester examinations that mixed multiple-choice and open-response questions. All examinations were written by the lead instructor. Conceptual learning was measured with the CSEM which was administered as a pretest and post-test in the laboratory [4]. Test averages and the normalized gain on the CSEM are presented in Fig. 1.

The course was presented in the above format for all of the 20 semesters studied. Over the 20 semesters, the class enrollment grew from an average yearly enrollment of 232 to 398. The number of students per laboratory section was constant. The course was generally well liked by students, with the lead instructor and the course itself receiving high student evaluations relative to other required science classes.

The course studied was primarily designed to be a high quality educational experience and was identified as one of

the important factors in the 10-fold growth in physics majors graduated by the physics department since its initial revision. For more discussion of the course and its role in the recruitment of physics majors see the discussion in Stewart, Oliver, and Stewart [60]. The course was also explicitly designed to be a highly stable, well-controlled environment in which to conduct physics education research (PER). This study seeks to understand the effect of variations in the course on student regulation of out-of-class TOT; these effects are subtle and would often be unmeasurable against the background of variation in a university physics course where instructors change every few semesters bringing varying levels of PER knowledge, teaching experience, and motivational skill—where laboratory experiences are often not related to lecture experiences and TAs are poorly or unevenly trained in the pedagogical goals of the laboratory sessions—and where changes in instructors produce often dramatic changes in homework assignments and examination difficulty. All of

these random factors present in most university courses have been either controlled or characterized in the course studied. All course materials were constructed specially for the course, so variation in the textbook can be characterized. The course features carefully scripted lectures and laboratory introduction talks all captured digitally. The solutions to all assignments including laboratories with the desired presentation are provided to all course staff. For the period studied, the same lead instructor presented the lecture, designed the homework assignments, quizzes, and examinations, oversaw TA training for the laboratory, and presented the first laboratory session which was also used to train new TAs. All course materials and assignments were captured in an electronically analyzable format and archived for the period studied, allowing the difficulty of the homework to be monitored. These efforts made for an exceptionally stably presented course except for planned revisions and the natural variation in assignments which were quantitatively characterized.

B. Survey design and validation

Students' out-of-class behaviors were measured with two surveys asking a variety of questions about topics ranging from the amount of time spent working homework sets to how thoroughly the student read the textbook. The survey questions were constructed first by examining required assignments for the class such as homework, reading, and examinations. Optional resources provided by the class that could be used, such as office hours and practice tests were also included as were observed student behaviors such as the taking of lecture notes. Behaviors thought to improve learning such as reading the textbook before attending lecture were also captured. Out-of-class behaviors not initially captured were identified through student interviews, student journals, and open response questions given in preliminary survey instruments. Only the results of a questions pertaining to general out-of-class time use were used for this study. A general analysis of the responses for semesters 1 to 8 was presented in Stewart, Stewart, and Taylor [3].

The surveys underwent an extensive construction and validation process during the two years before data collection began for this study. Preliminary versions of the surveys were constructed by examining course resources and policy including reading and homework policy, testing policy, and resources made available to students such as practice tests and office hours. This produced instruments that collected itemized estimates of the time invested in a variety of out-of-class academic activities. Sixty students were asked to keep journals detailing their out-of-class activities and the time invested in those activities. The journals were collected and examined to identify out-of-class academic behaviors not represented in the preliminary surveys. Robinson had previously demonstrated that time diaries were effective in gathering accurate time self-reports

[61] and time diaries have since been used in a number of studies [62,63]. The survey instruments were modified to include the additional behaviors and then applied to the sixty participating students. Forty-three of the students both returned their journal and completed the survey instrument. The average absolute value of the percent difference between the total of the itemized self-reported TOT responses and the journals was 18%. Approximately the same number of students overestimated out-of-class time (19) as those who underestimated (20), with 4 students estimating time correctly. The average percent difference between the journal and survey total was -1.3%. The survey instrument and the journals were organized differently so this small difference for time estimation could not be the result of recalling the journal entries. The students were also asked for an estimate of their total out-of-class time in the previous week at the end of the in-semester examination given at the end of the journaling experiment. The responses given to this single question were much less predictive of the journal totals with an average absolute percent difference of 38%. We theorize that the detailed information asked for in the surveys caused the students to consider their out-of-class actions more carefully than a single question given after an examination. Discussions with students supported this hypothesis. This mechanism for improving the accuracy of student self-reports was also utilized by Brint and Cantwell [64]. The approximately 40% discrepancy between journals and single-question time self-reports is very consistent with the finding of Steinbrickner and Steinbrickner [63], who found a correlation of 0.72 between self-reports and journals and an overestimation of time spent of 0.4 h per hour reported; strong correlations between self-reports and journals was also found for occupational time [61].

The full surveys used were constructed from this preliminary survey by adding additional questions, not used in this study, that collected the number of times students performed certain activities such as working practice tests and when and how thoroughly they performed certain activities such as reading the textbook. The completed surveys were tested in the spring and fall semesters before data collection for this study began. A subset of students were asked to discuss their understanding of the survey questions. The piloted surveys contained some open-response questions about study behavior. Responses to open questions in the surveys were analyzed and the instruments refined. The final instruments were given without substantial modification for the 20 semesters included in this study. The itemized time use questions were modified when an element was added to the class such as the self-testing tool introduced in semester 17. Mathematical reliability and construct validity analysis were not appropriate because the surveys collected lists of disparate actions so each answer was expected to vary independently.

While the above comparison of journal results suggest that the self-reported times were reasonably accurate and that students neither preferentially under nor overestimate time, it should be emphasized that what was actually measured was the student's impression of his or her time investment. The accuracy of this impression may vary from student to student. This variation makes comparisons of the same student at different times (assuming a student has the same pattern of misestimation) or comparisons of averages over random samples of students (those taken over the semester or academic year) more reliable than analysis that compares two different students.

Various measures support the validity of the survey instruments. The comparison of journal results and survey results give evidence of convergent validity. We will find the time investment reported was consistent with other published measures. Further, the analysis performed was consistent even when a primary threat to validity, student over reporting of time investment, was removed by scaling data taken at two time points by the average. To move beyond this level of support for the validity of the self-reported time is possible, but would require capturing student behavior in a controlled environment possibly by asking that they work homework and prepare for examinations in a facility where their actions could be recorded. This was not feasible for this study.

Survey development and application conformed to best practice identified in Kuh's work on self-reported data [65]. The surveys collected factual information known to the student about the recent past. The surveys were administered with strong guarantees of anonymity and contained simply worded questions asking for information that should not have discomforted the student. Laing *et al.* reported a 10% error rate in factual responses using student self-reports [66].

C. Data collection

The two surveys were given in the laboratory in the weeks following the first and third in-semester examination. The surveys were optional and students were told that the surveys would not be examined until the final course grades were submitted and that the lead instructor would only receive summary statistics. For the analysis which follows, only students who completed all questions on both surveys, who completed both the CSEM pretest and post-test, and who completed the class were included. The students included in the analysis were a somewhat different population from the full population of the class; they have somewhat higher attendance and test scores than the full population. For a more complete analysis of the two populations, see Stewart, Stewart, and Taylor [3].

D. Survey questions

The two surveys asked students a variety of questions about the kinds of out-of-class actions taken to prepare for

the class, how thoroughly the actions were performed, when they were performed, and how much time was spent performing them. This study will focus only on the responses to a small subset of the questions which asked about the time spent reading, working homework, and studying for the in-semester examinations. Both surveys asked how much time was spent working an average homework set. These responses were used with the number of homework sets collected to compute an average homework time per week. Survey 2 contained a question asking how much time was spent reading the course textbook per week, excluding examination study time, which was used as the reading time per week. Reading time per week was added to homework time per week to yield nonexam time per week, the average time spent out of class during weeks not containing an examination. Both surveys contained a question asking for the time spent studying for the in-semester examinations. Survey 2 also contained a question that asked the student to itemize his or her examination study time and then asked for the total study time again. The three questions were averaged to yield the exam study time, the time spent preparing for one of the in-semester examinations. As part of the itemization of examination study activities, the students were asked for the amount of time spent reading while studying for the examination, the exam study reading time. The exam study reading time was multiplied by the number of examinations, divided by the number of weeks in the semester, and then added to the reading time per week to yield the total reading time per week, the average time spent per week reading the course textbook. The exam study time was also averaged over the semester by multiplying by the number of examinations and dividing by the weeks in the semester and added to the nonexam study time to yield the total time per week, the average time per week spent on out-of-class activities during the semester. The difference between the examination preparation time for tests 1 and 3 was calculated as the change in exam study time while the difference in the time invested in a homework set from test 1 to 3 was calculation as the change in homework set time. The TOT variables are summarized in Table I.

E. Reading and homework length

In addition to the results of the surveys, test averages, and CSEM normalized gains, two additional variables will be used to characterize the course, the lengths of the course textbook and the homework assignments. The length of the course textbook, the reading length, was defined as the number of characters in the LaTeX file used to generate the textbook.

The length of the homework assignments was more difficult to characterize because a problem that can be stated in a few words is often more challenging than a problem that requires many words to describe. The length of a homework problem was characterized utilizing a rubric

TABLE I. Summary of time-on-task variables.

Homework set time	The average time spent working a homework assignment.
Change in homework set time	The difference in the time invested in a homework assignment (set) between the first and second survey.
Homework time per week	The time spent working homework per week averaged over the semester.
Exam study time	The average time spent preparing for an examination.
Change in exam study time	The difference in the time invested in examination preparation between the first and second survey.
Reading time per week	The average time spent reading the textbook per week in weeks not containing an examination.
Total reading time per week	The average time spent reading per week, including reading done to prepare for an examination, averaged over the semester.
Nonexam time per week	The total out-of-class time for weeks not containing an examination.
Total time per week	The average time per week including the time spent preparing for examinations.

for dividing the solution to a physics problem into fundamental steps developed as part of a characterization system for physics problems (DUE-0535928). The rubric called for the subdivision of the text of the solution into fragments such that each retained independent meaning but could not be further subdivided without loss of meaning. These fragments were the sentence groups, clauses, and phrases within the solution that conveyed a single idea. Application of the rubric to examples in popular textbooks produced an agreement of 94% between multiple raters. The length of a homework problem was characterized by the number of these indivisible fragments, homework steps, required for its solution.

Not all steps are of equal difficulty or require equal time to perform; however, no explicit effort to select more or less challenging problems was reported by the lead instructor who constructed all class assignments. As such, the step difficulty should vary randomly and the time required for the completion of the assignment with the same level of diligence should be proportional to the number of steps.

F. Course revision

The lead lecturer for the 20 semesters studied took over the course from its original designer in Fall 2001. The Fall 2001 and Spring 2002 semesters were used to convert a set of course notes into the original version of the textbook. The course was presented in the structure described above for 20 semesters from Fall 2002 (semester 1) to Spring 2012 (semester 20).

While the overall model was maintained, the course did undergo modifications during the 10 years studied. Homework assignments, examinations, and quizzes were changed each semester producing fluctuations in homework length, test average, and normalized gain. Partially online homework assignments were introduced in semester 9 where some multiple-choice questions were answered online while the remaining problems were submitted on paper. A computerized self-testing tool that allowed students to generate small 75% conceptual, 25% quantitative online tests was introduced in semester 17. This system contributed to the increase in exam study time in semesters

17 to 20. The College of Engineering began introduction of a Freshman Engineering Program (FEP) in semester 11 to promote retention. One feature of this program was that all engineering students were encouraged to complete their core mathematics and science courses on the same schedule. This had the effect of enhancing the fluctuations in the size of the spring and fall classes as seen in Fig. 1. The FEP did not produce a substantial change in the class standing of the students in the class. The average standing changed from 2.3 before the FEP to 2.2 after its implementation; standing was calculated with freshman as 1, sophomores as 2, juniors as 3, and seniors as 4.

Two major course revisions were implemented during the study; one beginning in semester 5 and one in semester 9. The course textbook, some of the laboratory activities, and the course timing were revised after semester 4. The textbook was revised to add some additional conceptual material to fully support the hands-on activities done in the laboratory, to increase the amount of calculus used, and to fully support the concept of integration as the limit of a sum of small elements. One day was removed from the lecture schedule to provide a day to discuss the final examination. The contraction of the lecture schedule necessitated the reshuffling of some laboratory activities; all laboratory activities were examined and the least effective replaced. These changes resulted in an increase in the length of the textbook, an increase in normalized gain on the CSEM, and a decrease in test averages, as shown in Fig. 1. The lead instructor attributed the decrease in test averages to the increased coverage and increased use of calculus; the increase in conceptual gain may have resulted from the additional conceptual material in the reading or from the modified laboratory activities. Preliminary analysis of the TOT data for semesters 5 to 8 suggested that the increased reading commitment and coverage, and the increased mathematical complexity of some of the additional material, was outside of the comfort level of the student population. As such, starting in semester 9, some of the material which had been added in semester 5 was removed. All conceptual material was retained and the laboratory activities were not modified.

All modifications were carried out to meet instructional goals in a continuous attempt to improve the course. These modifications produced fluctuations in the course difficulty and workload that served as a natural experiment which can be used to explore student regulation of time use.

IV. RESULTS

During the 20 semesters studied, 3116 students completed the class. Of these, the 1676 students who completed all four in-semester examinations, the CSEM pretest and post-test, and all questions on both surveys were included in this study. By semester averages of the number of students included in the study, N , the test average on the first two in-semester examinations, the total homework length for the semester in homework steps, the length of the textbook in characters, the normalized gain on the CSEM, and the nonexam time per week, the exam study time, and the total time per week are presented in Fig. 1. These plots show the planned changes in the length of the textbook, a decrease in test average with the revision beginning in semester 5, the increase in normalized gain with this revision, and the strong spring-fall oscillation in class size beginning in semester 12 and a weaker oscillation in test average. The graphs also show relatively stable test averages and normalized gains from semester 5 onward supporting that the efforts to make the course a stable research environment were successful. The homework length showed a pattern of decrease until semester 10 and then a pattern of increase from semesters 11 to 20. These changes were unplanned; the increasing homework length later in the study may have resulted from a growing pool of high quality, multiple-choice problems written specifically for the class.

The average nonexam time per week was quite stable with 15 semesters in the half-hour range of 4.1 to 4.6 h, a total range of 1.6 h, and mean of 4.3 ± 0.3 h. The average exam study time shows more variation with a range of 4.9 h and mean of 7.3 ± 1.2 h. The study time increased as the material was made more difficult in semesters 5 to 8, decreased slightly with the revision in semester 9, and then increased as the homework was made longer and the self-testing tool was released. Table II presents the correlation r with the student's score on the second in-semester examination, the average of the first two in-semester evaluations, and the average of all four in-semester examinations. Three correlations are presented because students may moderate their behaviors either primarily in reaction to the most recent examination or their overall examination average. The overall test average shows the effect of the student's behavior over the semester. While the second survey was given after the third in-semester examination, the students did not know the score on this examination during the preparation for the third examination. Table II uses three different aggregation methods. The overall entries were calculated by aggregating all data before calculating r ; the semester entries were

TABLE II. Summary of the correlation r of time on task with the score on the second in-semester examination, the average on the first two in-semester evaluations, and the overall in-semester evaluation average. The mean (M) and standard deviation (SD) are also reported.

	$M \pm SD$ (hours)	Test 2 r	Test 1-2 r	All tests r
Overall ($N = 1676$)				
Test 2 average	75.1 ± 17.5			
Test 1 and 2 average	76.8 ± 14.1			
Homework time	2.5 ± 1.2	-0.05	-0.07 ^b	-0.05 ^a
Reading time	1.8 ± 1.4	-0.00	-0.01	+0.00
Nonexam time	4.3 ± 2.1	-0.03	-0.05	-0.03
Exam study time	7.6 ± 4.8	-0.16 ^c	-0.20 ^c	-0.15 ^c
Total time	6.5 ± 2.9	-0.10 ^c	-0.13 ^c	-0.09 ^c
Semester ($N = 20$)				
Test 2 average	75.6 ± 4.9			
Test 1 and 2 average	77.0 ± 3.3			
Homework time	2.6 ± 0.2	+0.11	+0.27	+0.25
Reading time	1.8 ± 0.2	-0.11	-0.03	-0.18
Nonexam time	4.4 ± 0.3	-0.00	+0.15	+0.04
Exam study time	7.3 ± 1.3	-0.57 ^b	-0.62 ^b	-0.55 ^a
Total time	6.4 ± 0.5	-0.41	-0.35	-0.37
Academic year ($N = 10$)				
Test 2 average	76.0 ± 3.6			
Test 1 and 2 average	77.5 ± 2.4			
Homework time	2.6 ± 0.2	+0.47	+0.59	+0.45
Reading time	1.8 ± 0.2	+0.01	-0.17	-0.46
Nonexam time	4.3 ± 0.3	+0.31	+0.27	+0.00
Exam study time	7.3 ± 1.2	-0.69 ^a	-0.84 ^b	-0.76 ^a
Total time	6.4 ± 0.5	-0.33	-0.47	-0.57

^a $p < 0.05$.

^b $p < 0.01$.

^c $p < 0.001$.

calculated by first calculating an average for the semester and then calculating r for the semester averages. The academic year entries were calculated by aggregating all students for an academic year, calculating the average values, and then computing r for these averages.

Aggregating all participants, Table II shows a significant but negative correlation between examination score and the number of hours spent in examination preparation. At first counterintuitive, this indicates that students who were performing more weakly on the examinations reported investing more time in examination preparation; students who are not doing well in a class invest more time in the class. Not all students have equal facility with the material presented in a physics class; those not receiving the test grades desired invest more effort. Small, nonsignificant, correlations were observed for the time spent in other out-of-class activities (nonexam time); some correlations with the amount of time spent working homework were significant, but much smaller than those with exam study time.

Semester and academic year averages show a strengthened correlation with exam study time and exposed a

relatively large positive correlation with homework time; semesters or academic years where students on average invested more time working homework produced higher test scores. These positive correlations with working homework were balanced with negative correlations with reading during nonexam weeks; reading time and homework time combine to form nonexam time, which consistently showed very small correlations. The time students invest outside of preparing for examinations had very little correlation with the scores on those examinations. Students were not moderating their time investment in class activities in response to their test average except by increasing the amount of time spent in exam preparation.

A. Regulation of examination preparation time and test average

The most important feedback provided to the students in the class studied was the scores on the four in-semester examinations. Examination grades accounted for approximately 70% of the total grade in the course. Both surveys contained questions asking how much time was invested in preparing for the most recent examination and, therefore, the change in examination preparation time, the change in exam study time, could be calculated between the first and third in-semester examinations. The correlation between test performance and the change in preparation time is presented in Table III. This correlation is a measure of how much students adjust their out-of-class behavior in reaction to the stimulus of examination grades. All students were expected to spend more time preparing for exam 3 which

covered somewhat more material than exam 1 and a topic, magnetism, that many students find challenging. Correlations determine whether students with weaker exam 2 scores or exam 1 and 2 averages increased their time investments more than stronger performing students. The change in exam study time is plotted against the score on the second examination averaged over the semester in Fig. 2 and averaged over the academic year in Fig. 3. The correlation of the test 2 score with the change in exam study time aggregating all data yielded $r = -0.12$ [$t(1674) = -4.91$, $p < 0.0001$], when aggregated by semester (Fig. 2) $r = -0.70$ [$t(18) = -4.14$, $p = 0.0006$], and when aggregated by academic year (Fig. 3) $r = -0.82$ [$t(8) = -4.01$, $p = 0.0039$]. The results were similar to those obtained with test 1 and 2 average as the independent variable as shown in Table III. Test 2 score explains 1% of the variance in change in exam study time overall, 49% by semester, and 67% by academic year. These results indicate that students do regulate their examination preparation time in response to their previous performance on examinations in a science class. The correlation between the time invested in studying for examinations and the score on the second examination changed dramatically for the different methods of averaging used in the three sections of Table III. The correlations were very similar for the test 1 and 2 average and the test 2 score. The test 2 score generally produced somewhat larger correlations (Table III). Figure 3 also presents the linear fit to the data, drawn as a dashed line, eliminating the first two academic years to test the influence of these points; the slope of the plotted line changed little removing these points.

TABLE III. Summary of the correlation of time on task and change of time on task with the score on the second in-semester examination and the average on the first two in-semester evaluations. The correlation is represented by r ; r_{frac} represents the correlation with the fractional change in the variable. CI represents the 95% confidence interval for the correlation. Change in exam study time and homework time is measured in hours; change in homework average and submission rate in percent.

Change in	Test 2					Test 1 and test 2 average				
	$M \pm SD$	r	CI	r_{frac}	CI	r	CI	r_{frac}	CI	
Overall ($N = 1676$)										
Exam study time	$+3.0 \pm 5.0$	-0.12^c	-0.17, -0.07	-0.09^c	-0.14, -0.04	-0.12^c	-0.17, -0.07	-0.07^b	-0.12, -0.02	
Homework set time	-0.2 ± 1.0	+0.03	-0.02, +0.08	+0.00	-0.05, +0.05	+0.04	-0.02, +0.09	+0.03	-0.02, +0.08	
Homework submission	-1.1 ± 13.3	-0.06^a	-0.11, -0.01	-0.07^b	-0.12, -0.01	-0.04	-0.09, +0.01	-0.04	-0.10, +0.02	
Homework average	$+0.1 \pm 15.3$	-0.16^c	-0.22, -0.11	-0.15^c	-0.21, -0.09	-0.13^c	-0.18, -0.08	-0.11^c	-0.17, -0.05	
Semester ($N = 20$)										
Exam study time	$+2.7 \pm 1.2$	-0.70^c	-0.87, -0.37	-0.63^b	-0.84, -0.26	-0.50^a	-0.77, -0.07	-0.34	-0.68, +0.12	
Homework set time	-0.2 ± 0.2	-0.15	-0.55, +0.32	-0.20	-0.59, +0.26	+0.03	-0.41, +0.47	+0.00	-0.44, +0.44	
Homework submission	-0.9 ± 2.8	-0.12	-0.53, +0.34	-0.11	-0.53, +0.35	-0.09	-0.51, +0.37	-0.08	-0.51, +0.37	
Homework average	-0.0 ± 5.1	-0.26	-0.63, +0.21	-0.26	-0.63, +0.20	-0.12	-0.54, +0.34	-0.13	-0.54, +0.33	
Academic Year ($N = 10$)										
Exam study time	$+2.8 \pm 1.2$	-0.82^b	-0.96, -0.39	-0.81^b	-0.95, -0.37	-0.77^b	-0.94, -0.27	-0.63	-0.90, +0.00	
Homework set time	-0.2 ± 0.2	-0.31	-0.79, +0.40	-0.36	-0.81, +0.35	-0.03	-0.65, +0.61	-0.07	-0.67, +0.58	
Homework submission	-1.0 ± 1.5	+0.04	-0.60, +0.66	+0.07	-0.59, +0.67	+0.26	-0.45, +0.76	+0.24	-0.46, +0.75	
Homework average	$+0.1 \pm 3.5$	-0.10	-0.69, +0.56	-0.10	-0.69, +0.57	+0.01	-0.62, +0.64	+0.00	-0.63, +0.63	

^a $p < 0.05$.
^b $p < 0.01$.
^c $p < 0.001$.

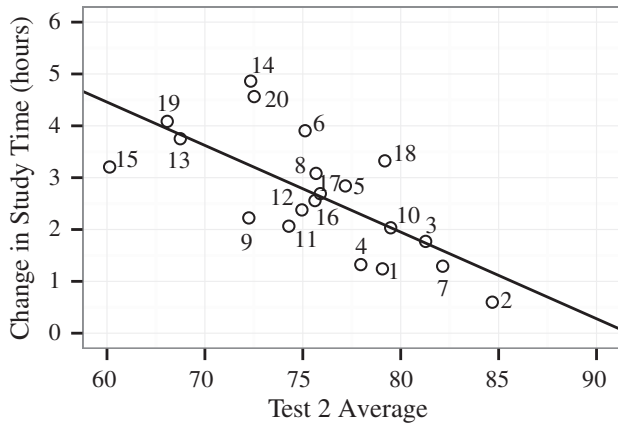


FIG. 2. Change in exam study time vs test 2 average by semester.

Table III reports both a significance level under the assumption of normality of the underlying distribution and the 95% confidence interval for all correlations. The overall aggregated test 2 average distribution was skewed with test average bounded above by 100%. To investigate the influence of deviation from normality, bootstrapping was used to draw 2000 subdistributions to establish the confidence intervals for the overall data. The confidence intervals support the conclusions drawn from the significance tests. The reported confidence intervals for the semester and academic year values were taken from the confidence interval for r assuming normality; an assumption that these data obey more strongly because of their more restricted standard deviations.

The possible effect of differentially accurate self-reporting of time use was also investigated. Students reporting 1 h of exam preparation time may be generally more accurate than those reporting 10 h. To investigate this, the self-reported times were converted to normalized rank

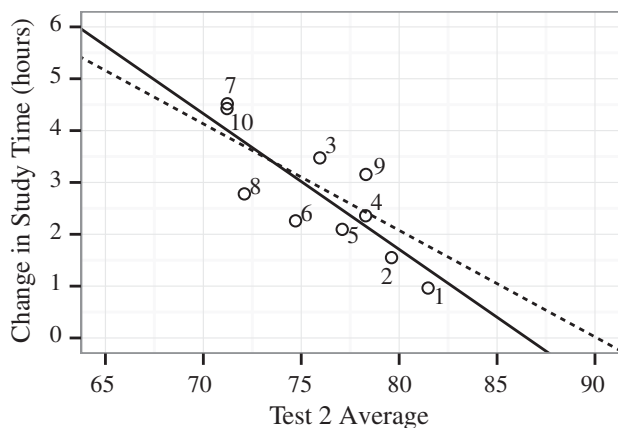


FIG. 3. Change in exam study time vs test 2 average by academic year. The solid line is the regression line using all 10 academic years; the dashed line eliminates semesters 1 and 2 to investigate the influence of these points.

data by replacing each data point with its probability calculated from the cumulative distribution function of the data. The correlation of test 2 score with the normalized change in exam study time was -0.14 ($p < 0.0001$). In general, the correlations with the normalized rank data were very little changed from the correlations reported in Table III.

With both the random changes to homework and tests and planned changes to reading and class policy as detailed in Sec. III F, using a statistical analysis method that treated the participants as nested within the semesters was appropriate. Both multiple-linear regression coding the semesters as dichotomous variables and hierarchical linear modelling treating the semesters as random effects were performed. The regression coefficients relating the time variables to performance were very similar for simple linear regressions without including the semesters as variables, multiple-linear regressions including the semester as a categorical variable, and hierarchical linear models treating the semester as a random effect. As such, the correlation coefficients, which are directly related to the simple linear regression coefficients, are reported in this paper because of their familiarity and their natural measure of effect size.

B. Fractional change in exam study time

The change in exam study time may be influenced by the overall negative correlation with exam study time; students who invest more time on average in examination preparation may also need to change their time investment more to have the same effect on test grade as students investing less time. To separate the effect of regulation from the correlation with exam study time, the change in exam study time was scaled by dividing by the average exam study time reported on the two surveys to form the fractional change in exam study time. Correlations with this variable are reported as r_{frac} in Table III. The fractional change in exam study time has the additional desirable property of correcting for misreporting of study time; if a student consistently inflates or underestimates his or her time investment, this misestimation will cancel out of the fractional change in exam study time. The fractional correlations were somewhat smaller, but still significant. The observations made for the correlations with change in exam study time also hold for the fractional change. The pattern of increasing correlation as the data was aggregated was repeated for the fractional change in exam study time giving evidence that the change was not an artifact of either the overall correlation with exam study time or a systematic pattern of misreporting time investment. The robustness of these correlations when scaled provides further evidence that the time reports were systematically related to student time investment.

C. Effect of increased aggregation

A number of analyses were used to explore the effect of increasing correlation with higher levels of aggregation.

Examination of Fig. 2 shows that fall semesters (odd numbers) cluster toward lower test averages and the first four semesters toward higher test averages. Spring semesters are generally larger and the class became larger with time. By plotting semester averages, the set of students in a semester is replaced with a single average, and thus the averaging weighs a student in a smaller semester more heavily than one in a larger semester. To determine if this was the source of the increasing correlation, larger semesters were sampled if they contained over 50 students before the data set was aggregated. This increased the aggregated correlation to -0.17^c , higher, but not close to the semester or academic year correlations.

The small size of the semester and academic years data sets could allow some semesters or years to exert unrepresentative influence on the correlation. This was investigated by sampling the distributions using an algorithm related to bootstrapping. If 2000 samples of 10 semesters are drawn from the 20 semester data set, the average correlation of test 2 score with change in exam study time was $r = -0.70$ with a 95% confidence interval of $(-0.88, -0.42)$. If 2000 samples of 5 years are drawn from the 10 academic years, the mean correlation of test 2 score and change in exam study time was $r = -0.78$ with a 95% confidence interval of $(-0.98, -0.27)$.

With the correlation relatively unchanged in both of the above analyses, the increased correlation with the semester average appears not to be an artifact of the statistical analysis. Individual student's behavioral regulation due to examination scores was stronger when the test average for the whole class was lower than the regulation when the student's score alone was lower. In semesters when the tests were actually more challenging (measured by class test average), there was more general behavioral regulation of examination preparation time.

D. Homework preparation behavioral regulation

A student can regulate his or her homework preparation behavior by investing more time in individual homework sets, by changing the rate at which homework sets are submitted, or by changing how well the sets are completed measured by the grade on the homework set. Both surveys asked for the number of hours spent working a homework set allowing the calculation of the difference in time invested in a homework set between test 1 and test 3, the change in homework set time as shown in Table III. To examine the change in homework submission rate and homework average, the difference in the rate at which homework sets were submitted and the difference in the average score on the homework sets was compared for the one month period from examination 1 to 2 and the one month period between examination 2 and 3 and the correlation with test average calculated as presented in Table III. Seven homework assignments were collected in each one-month period. Fractional versions of each variable

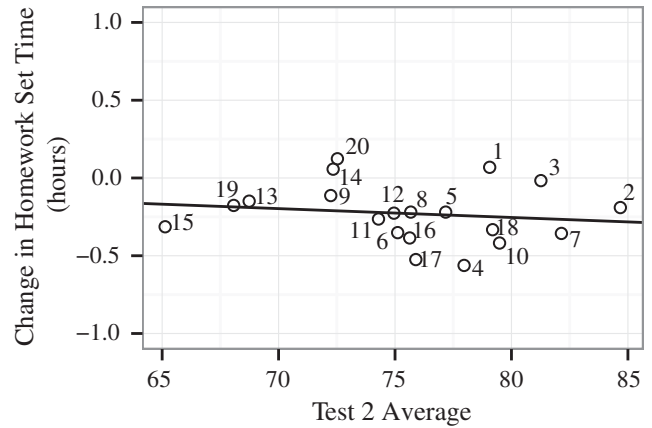


FIG. 4. Change in homework set time vs test 2 average by semester.

are also presented and were calculated by dividing the change in the quantity by the average of the two quantities used to compute the change.

Aggregating all students, Table III shows a very weak correlation between test 2 score and the time invested in a homework set, a weak but significant negative correlation with homework submission rate, and a stronger negative correlation with the change in homework average. The correlation with the change in average was stronger than the correlation with the change in study time. As the data was aggregated by semester and academic year, change in exam study time emerged as, by far, the variable most strongly correlated with test performance. The homework results shown in Table II, Table III, and Fig. 4 show weaker regulation of homework time investment than was evident in the regulation of examination preparation behavior; students reacted to changes in test performance with more studying, but not with changes in homework preparation time.

The results also showed a significant correlation between an increase in homework average between exams 2 and 3 and the score on exam 2 or the average of exam 1 and 2. This increased average was not accompanied by an increase in time invested or in the number of assignments submitted and, therefore, must have resulted in time neutral changes in homework behavior. These changes may have been educationally beneficial optimization of homework preparation to maximize learning or educationally wasteful inappropriate activities such as copying or overreliance on group answers. The effect of these changes can be investigated by regressing the change in homework average between test 2 and 3 on the change in test average between test 2 and 3 producing a significant linear model ($p < 0.0001$), but yielding a regression coefficient for the slope of only $\beta_1 = 0.14$. As such, a 10% increase in homework average yielded only a 1.4% increase in test average, suggesting that the changes in homework preparation behavior were not effective in producing improved test performance. Changes in behavior that increase

TABLE IV. Summary of the correlation of the time on task and the change of time on task with the average on the first two tests and the score on the second test separated by the grade received in the class.

	Grade A ($N = 760$)		Grade B ($N = 530$)		Grade C ($N = 161$)		Grade D ($N = 31$)	
Test 2 average								
	r	r_{frac}	r	r_{frac}	r	r_{frac}	r	r_{frac}
Change in exam study time	-0.21 ^c	-0.13 ^c	-0.22 ^c	-0.20 ^c	-0.02	+0.01	+0.05	+0.07
Change in homework set time	-0.01	-0.09 ^a	-0.01	-0.05	+0.00	+0.03	-0.11	-0.21
Change in homework Submission	-0.13 ^b	-0.06	-0.13 ^b	-0.14 ^b	-0.12	-0.07	-0.31	-0.27
Change in homework Average	-0.20 ^c	-0.13 ^c	-0.20 ^c	-0.20 ^c	-0.15	-0.10	-0.07	-0.03
Nonexam time per week	-0.01		-0.01		-0.14		-0.07	
Exam study time per exam	-0.11 ^b		-0.11 ^b		-0.08		-0.19	
Total time per week	-0.07		-0.07		-0.13		-0.12	
Test 1 and 2 average								
	r	r_{frac}	r	r_{frac}	r	r_{frac}	r	r_{frac}
Change in exam study time	-0.23 ^c	-0.16 ^c	-0.23 ^c	-0.16 ^c	-0.02	+0.00	-0.11	-0.04
Change in homework set time	+0.04	-0.07	+0.04	-0.01	-0.06	+0.03	-0.11	-0.17
Change in homework submission	-0.09 ^a	-0.07 ^a	-0.09 ^a	-0.11 ^a	-0.17 ^a	-0.13	-0.14	-0.13
Change in homework Average	-0.14 ^b	-0.12 ^b	-0.14 ^b	-0.14 ^b	-0.20 ^a	-0.12	+0.11	+0.17
Nonexam time per week	-0.06		-0.06		-0.25 ^b		-0.12	
Exam study time per exam	-0.21 ^c		-0.21 ^c		-0.20 ^a		-0.27	
Total time per week	-0.16 ^c		-0.16 ^c		-0.26 ^c		-0.20	

^a $p < 0.05$.^b $p < 0.01$.^c $p < 0.001$.

homework scores without increasing understanding (copying, overreliance on group work) are very time efficient explaining the failure to observe a change in homework preparation time. As was found with the change in exam study time, the fractional correlations were very similar to those found without scaling the variables.

E. Behavioral regulation by final class grade

The extent to which the above results varied with students achieving different outcomes in the class was also investigated by separating the students by their final class grade and repeating the analysis. Final grade data was not available for semesters 1 to 3. The pattern of weak behavioral regulation of the time spent working on homework was evident at all final grade levels as shown in Table IV. Regulation of examination preparation time was nearly identical for students earning an “A” or “B” in the class with stronger correlations with test average than those calculated for the class as a whole. The correlation of the change in exam study time with test 2 score changed dramatically for students earning a “C” or “D” in the class; these students showed very little behavioral regulation of their time investment in examination preparation in response to their performance on the first two examinations or the most recent examination. A t test showed that students receiving an “A” or “B” had statistically significantly different change in exam study time [$t(238) = -2.17$, $p = 0.0311$] than lower performing students but that the change in homework set time was not significant.

F. Behavioral regulation and reading length

The total time per week spent reading the course textbook is plotted against the length of the course textbook in Fig. 5. The data are well fit by a linear function with $R^2 = 0.28$ [$F(1, 18) = 7.11$, $p = 0.0157$] also plotted in Fig. 5. This line, however, has an intercept that is substantially different from zero. Students cannot spend time reading the textbook if there is no textbook, and therefore some other nonlinear functional form that meets the requirement of a zero intercept must fit the data. The limited range of the data does not allow the resolution of the exact form of the required function. A search of the literature did not uncover

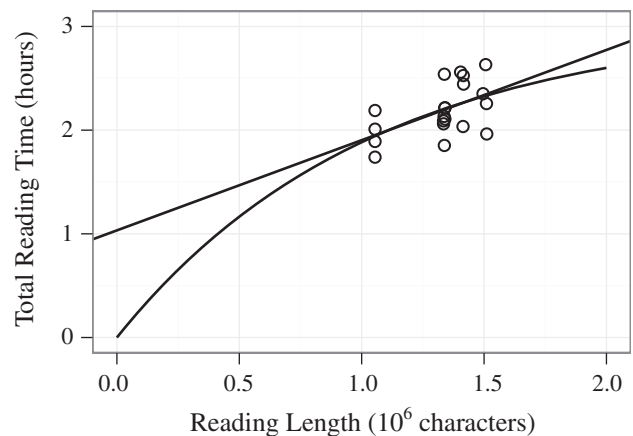


FIG. 5. The total time in hours spent reading per week plotted against the total length of the course textbook in millions of characters.

research identifying the functional form of the growth of time investment with assignment length. As such, we propose a function with the correct qualitative behavior, the increasing exponential function. The required function $t_r(\ell)$, where t_r is the reading time in hours and ℓ is the length of the course textbook in millions of characters, must be zero at $\ell = 0$, and, because students cannot invest an unlimited time in a single course, should approach some maximum value $t_{r,\max}$. The increasing exponential function $t_r(\ell) = t_{r,\max}(1 - e^{-\ell/\tau_R})$ has the required behavior where τ_R is a parameter characterizing how quickly $t_{r,\max}$ is approached. Fitting this function to the data yielded $t_{r,\max} = 3.0$ h and $\tau_R = 1.0 \times 10^6$ characters. The function $t_r(\ell)$ is also plotted in Fig. 5.

The students increased their time spent reading the textbook but did not invest additional time in proportion to the changes in the textbook. To the extent that the time spent reading a passage is a measure of the care spent reading it, students read less carefully as the text increased in length. The reading commitment for the course ranges from 1.1 to 1.5×10^6 characters or from $1.0\tau_R$ to $1.5\tau_R$; therefore, some additional reading commitment could be extracted from the students, but most of the time available was already committed. At $1.5\tau_R$, t_r is 77% of $t_{r,\max}$.

G. Behavioral regulation and homework length

The homework assignments for the class were changed each semester and were composed of previously used problems with a small number of new problems each semester. All problems were written specifically for the course studied. The assignments were assembled by the lead instructor with the intention of giving approximately the same amount of homework and homework of equivalent difficulty each semester; however, despite this intention, homework assignments showed a pattern of decreasing length from semesters 1 to 10 and then increasing length from 10 to 20.

The homework time per week is plotted against the total number of steps required to solve the homework in the semester in Fig. 6. A linear regression yields a line with $R^2 = 0.03$ [$F(1, 18) = 0.55$, $p = 0.4681$], also plotted on Fig. 6. The small R^2 of this line results primarily from the small slope of the line, not a poor fit of the data.

As was found with the reading time, the intercept of the regression line was far from zero. If no homework was assigned, students would not invest time in the homework; therefore, the correct function must have a zero intercept. As before, the range of the data is insufficient to determine the function fitting the data, but an increasing exponential has the required qualitative behavior. The data were fit to the function $t_h(s) = t_{h,\max}(1 - e^{-s/\tau_H})$, where $t_{h,\max}$ is a constant representing the maximum homework time per week, s is the number of homework steps per week, and τ_H is a constant controlling how quickly $t_{h,\max}$ is approached. Fitting this function yielded

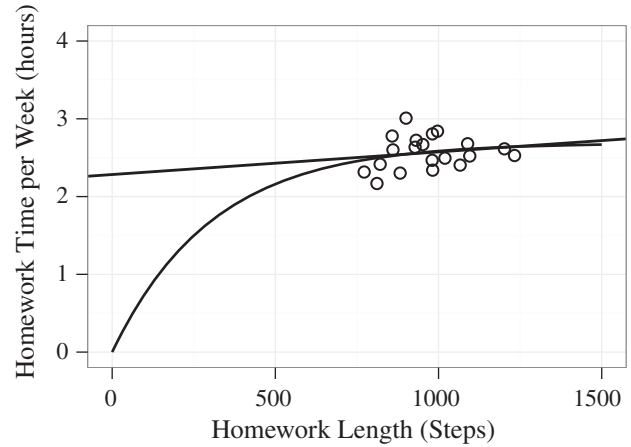


FIG. 6. The total time in hours spent working homework per week plotted against the total length of the homework assignments as measured by the number of steps required for their completion.

$t_{h,\max} = 2.7$ h and $\tau_H = 310$ steps per semester. This function is also plotted in Fig. 6.

Figure 6 shows little evidence that students modified the time invested in homework as more homework was assigned. This provides further support for the hypothesis that regulation of out-of-class behavior is directed almost exclusively to examination preparation. The average number of homework steps per semester ranges from 780 to 1230 or from $2.5\tau_H$ to $4.0\tau_H$. At $4.0\tau_H$, t_h is 98% of $t_{h,\max}$ and therefore for this student population the amount of homework time was very close to the maximum time they were willing to invest.

V. DISCUSSION

This study sought to determine if students in a physics class modify (regulate) their out-of-class academic behavior as a result of stimuli provided by the class in the form of examination scores and assignment (reading and homework) length. Is student time investment in a science class fixed or mutable? The students displayed a significant regulation of examination preparation time investment at the individual level in response to examination scores; this change in examination time investment remained significant when the time was scaled by the student's average total study time. While statistically significant, the correlations represent a small effect size. Functionally, the correlation of -0.12 between exam 2 score and the change in exam study time implies that a change in 1 standard deviation of test 2 score (18%) produced a change of 0.12 standard deviations of change in exam study time or 0.9 h. Correlations were fairly consistent when calculated with either the most recent test (exam 2) or the students' test average in the class before taking exam 3. Correlations with the fractional increase in TOT were also relatively consistent with the unscaled values indicating that students who were already

having to commit substantial time to the class increased their time commitment proportionally. The consistency between the fractional and absolute changes provides evidence that the significant relationship does not result from consistent misestimation of time investment.

Correlations with the overall measured variables presented in Table II are consistent with a body of research showing a weak relationship between the overall time investment and academic success [29]. The small negative correlation of homework time with test average does not support the body of research finding significant positive correlations with homework time investment [26], but supports more recent work that controlled for confounding variables such as school characteristics [27].

The changes in examination preparation time were examined aggregating the data by semester and academic year. The test 2 score explained 67% of the variance in the change in examination preparation time when the data were pooled by academic year. This suggests that class-level changes can generate substantially more behavioral modifications than those observed in individual students in response to their examination scores.

This study also sought to determine if behavioral regulation patterns differed by performance level. The regulation of examination preparation time was very consistent for students who earned an “A” or “B” in the class, but was significantly reduced for lower performing students; time regulation was dramatically different by performance level. This supports previous work that shows effort regulation as an important variable in academic performance [47].

Regulation of time invested in reading the course textbook was also investigated. The time invested in reading the textbook increased as the length of the reading assignments increased, but not proportionally to the change in reading length; at the minimum assigned reading length, a 10% increase in the length of the textbook produced only a 5% increase in the time spent reading.

Very little regulation of the time invested in homework was detected either in response to examination scores or homework assignment length; these students committed a fixed time to the homework regardless of external influences. This inflexibility can be seen in the strong saturation of homework time in Fig. 6 and the small correlations in Tables II and III. The inflexibility in time investment is consistent with the results of the National Survey for Student Engagement, which found little difference in time investment between freshmen and seniors [2].

While substantial out-of-class work was assigned in the course, the 4.3 h per week invested in weeks not containing an examination was less than the 8 h suggested for a four-credit class and less than the 6 h per week reported by Di Stefano [10] or the 5.3 h reported by Kortemeyer [11].

This work provides a more nuanced picture of student effort regulation and time management with differing

amounts of regulation directed toward various academic behaviors. Much more regulation of examination preparation time was observed than that of homework or reading time. More regulation was measured for large-scale, semester-level changes in the class than was observed in individual student performances.

VI. IMPLICATIONS FOR INSTRUCTION

Students demonstrated inflexibility in the amount of time they were willing to invest in the class outside of the time spent preparing for an examination. This implies that the time required to complete the assignments in a class in such a manner that the maximum learning occurs must be fit to this fixed time allotment; assignments must be “right sized.” If material does not fit into the fixed time investment, students attempt to master the material by investing more time in preparing for the examinations. It is unlikely that this one-time investment of additional effort results in the same deep, persistent learning that results from an integrated learning experience combining all of the experiences available in the learning activities not involved in examination preparation. As one designs changes in a science class, this study suggests that nonstudy time should be viewed as a fixed resource, while examination preparation time can be considered a weakly mutable variable.

The failure of students to substantially adjust their behavior to changes in the length of the homework suggest that instructors must make careful choices to increase the educational value of individual homework problems. This may include using problems that are more contextually rich, involve different modes of reasoning, or ask the students to explore different representations. The failure to modify time investment as assignment length increased may mean that available out-of-class time is a limitation on reformed educational designs that require increased out-of-class TOT.

The changes in homework time investment measured involved traditional qualitative and quantitative problem-based homework assignments. The inflexibility in time investment found may not extend to different types of take home assignments or homework regimens that mix types of assignments. For example, it is quite possible that if a take-home video analysis project had been assigned as part of the homework that students would have found additional time to invest in the project. More research would be required to determine the extent to which the inflexibility of time investment measured in this study extended to other take home assignment types.

VII. FUTURE WORK

This study investigated only the most broad categories of time use in a physics class; the detailed way students allocate time to different study behaviors and how that time allocation changes through the class could provide a finer grained picture of time management. Surveys measuring

time use could be combined with surveys asking questions about students' impressions of resource management and effort regulation to determine how students' beliefs about their student habits are related to measurable changes in their study behavior. Subscales from the Motivated Learning Strategies Questionnaire could be used for this purpose [33].

VIII. LIMITATIONS

This research was performed at a single institution and therefore may only represent the behavior patterns of students at that institution. To determine if the findings are general, similar studies would have to be done at other universities. Many factors influence student behavior in a science class; this work was done in as controlled and well understood a course environment as possible, but the conclusions could be influenced by uncontrolled factors. As Di Stefano notes, a student's response to a science class is complex [10]; the measurement presented collects only overall self-reported averages of student time use, a more detailed measurement might produce additional insights.

IX. CONCLUSION

Students regulated the amount of time invested in examination preparation in response to their examination scores; this regulation was more pronounced as the examination average for the class as a whole changed in comparison with the examination average of individual students. The time spent reading did not scale proportionally with the length of the reading assignment; therefore, as assignments become longer, less time is invested for a given length of reading assignment. There was no evidence of regulation of homework time investment; the time spent working homework did not change in response to either assignment length or examination average.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation as part of the evaluation of improved learning for the Physics Teacher Education Coalition, PHY-0108787, and through the TOPP: Taxonomy of Physics Problems, Improving Student Understanding in Introductory Physics Grant No. DUE-0535928.

-
- [1] A. W. Chickering and Z. F. Gamson, Seven principles for good practice in undergraduate education, *AAHE Bull.* **39**, 3 (1987).
 - [2] National Survey of Student Engagement, *NSSE 2012 U.S. Grand Frequencies* (Indiana University Center for Post-secondary Research, Bloomington, IN, 2012).
 - [3] J. Stewart, G. Stewart, and J. Taylor, Using time-on-task measurements to understand student performance in a physics class: A four-year study, *Phys. Rev. ST Phys. Educ. Res.* **8**, 010114 (2012).
 - [4] D. P. Maloney, T. L. O'Kuma, C. Hieggelke, and A. Van Huevelen, Surveying students' conceptual knowledge of electricity and magnetism, *Am. J. Phys.* **69**, S12 (2001).
 - [5] H. T. Hudson and D. Liberman, The combined effect of mathematics skills and formal operational reasoning on student performance in the general physics course, *Am. J. Phys.* **50**, 1117 (1982).
 - [6] D. Liberman and H. T. Hudson, Correlation between logical abilities and success in physics, *Am. J. Phys.* **47**, 784 (1979).
 - [7] I. A. Halloun and D. Hestenes, The initial knowledge state of college physics students, *Am. J. Phys.* **53**, 1043 (1985).
 - [8] D. E. Meltzer, The relationship between mathematics preparation and conceptual learning gains in physics: A possible, hidden variable, in diagnostic pretest scores, *Am. J. Phys.* **70**, 1259 (2002).
 - [9] 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).
 - [10] R. Di Stefano, Physics teaching and time management, *Phys. Teach.* **36**, 350 (1998).
 - [11] G. Kortemeyer, Gender differences in the use of an online homework system in an introductory physics course, *Phys. Rev. ST Phys. Educ. Res.* **5**, 010107 (2009).
 - [12] W. W. Welch and R. G. Bridgman, Physics achievement gains as a function of teaching duration, *School Sci. Math.* **68**, 449 (1968).
 - [13] L. Springer, M. S. Stanne, and S. S. Donovan, Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis, *Rev. Educ. Res.* **69**, 21 (1999).
 - [14] S. Kolari, C. Savander-Ranne, and E. L. Viskari, Do our engineering students spend enough time studying?, *Eur. J. Eng. Educ.* **31**, 499 (2006).
 - [15] S. Kolari, C. Savander-Ranne, and E. L. Viskari, Learning needs time and effort: A time-use study of engineering students, *Eur. J. Eng. Educ.* **33**, 483 (2008).
 - [16] W. Fakcharoenphol, J. W. Morpew, and J. Mestre, Judgments of physics problem difficulty among experts and novices, *Phys. Rev. ST Phys. Educ. Res.* **11**, 020128 (2015).
 - [17] J. W. Morpew, J. P. Mestre, B. H. Ross, and N. E. Strand, Do experts and novices direct attention differently in examining physics diagrams? A study of change detection using the flicker technique, *Phys. Rev. ST Phys. Educ. Res.* **11**, 020104 (2015).
 - [18] K. Miller, N. Lasry, B. Lukoff, J. Schell, and E. Mazur, Conceptual question response times in peer instruction classrooms, *Phys. Rev. ST Phys. Educ. Res.* **10**, 020113 (2014).

- [19] A. Renkl, Learning from worked-out examples: A study on individual differences, *Cogn. Sci.* **21**, 1 (1997).
- [20] M. Romero and E. Barberà, Quality of e-learners time and learning performance beyond quantitative time-on-task, *Int. Rev. Res. Open Dist. Learn.* **12**, 125 (2011).
- [21] B. Means, Y. Toyama, R. Murphy, M. Bakia, and K. Jones, *Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies* (U.S. Department of Education, Office of Planning, Evaluation, and Policy Development, Washington, D.C., 2009).
- [22] D. J. Palazzo, Y. J. Lee, R. Warnakulasooriya, and D. E. Pritchard, Patterns, correlates, and reduction of homework copying, *Phys. Rev. ST Phys. Educ. Res.* **6**, 010104 (2010).
- [23] 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).
- [24] A. Goldstein, Does homework help? A review of research, *Elem. Sch. J.* **60**, 212 (1960).
- [25] R. A. Paschal, T. Weinstein, and H. J. Walberg, The effects of homework on learning: A quantitative synthesis, *J. Educ. Res.* **78**, 97 (1984).
- [26] H. Cooper, J. C. Robinson, and E. A. Patall, Does homework improve academic achievement? A synthesis of research, 1987–2003, *Rev. Educ. Res.* **76**, 1 (2006).
- [27] U. Trautwein, The homework–achievement relation reconsidered: Differentiating homework time, homework frequency, and homework effort, *Learn. Instr.* **17**, 372 (2007).
- [28] R. De Jong, K. J. Westerhof, and B. P. M. Creemers, Homework and student math achievement in junior high schools, *Educ. Res. Eval.* **6**, 130 (2000).
- [29] E. A. Plant, K. A. Ericsson, L. Hill, and K. Asberg, Why study time does not predict grade point average across college students: Implications of deliberate practice for academic performance, *Contemp. Educ. Psychol.* **30**, 96 (2005).
- [30] J. E. Dowd, I. Araujo, and E. Mazur, Making sense of confusion: Relating performance, confidence, and self-efficacy to expressions of confusion in an introductory physics class, *Phys. Rev. ST Phys. Educ. Res.* **11**, 010107 (2015).
- [31] K. Miller, J. Schell, A. Ho, B. Lukoff, and E. Mazur, Response switching and self-efficacy in peer instruction classrooms, *Phys. Rev. ST Phys. Educ. Res.* **11**, 010104 (2015).
- [32] B. J. Zimmerman, Self-regulated learning and academic achievement: An overview, *Educ. Psychol.* **25**, 3 (1990).
- [33] P. R. Pintrich, D. A. F. Smith, T. García, and W. J. McKeachie, Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ), *Educ. Psychol. Meas.* **53**, 801 (1993).
- [34] A. L. Duckworth, C. Peterson, M. D. Matthews, and D. R. Kelly, Grit: Perseverance and passion for long-term goals, *J. Pers. Soc. Psychol.* **92**, 1087 (2007).
- [35] P. R. Pintrich, A conceptual framework for assessing motivation and self-regulated learning in college students, *Educ. Psychol. Rev.* **16**, 385 (2004).
- [36] A. Efklides, Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model, *Educ. Psychol.* **46**, 6 (2011).
- [37] A. Koriat, The relationships between monitoring, regulation and performance, *Learn. Instr.* **22**, 296 (2012).
- [38] B. J. Zimmerman and M. M. Pons, Development of a structured interview for assessing student use of self-regulated learning strategies, *Am. Educ. Res. J.* **23**, 614 (1986).
- [39] A. Bandura, Self-efficacy mechanism in human agency, *Am. Psychol.* **37**, 122 (1982).
- [40] D. H. Schunk, Self-efficacy perspective on achievement behavior, *Educ. Psychol.* **19**, 48 (1984).
- [41] B. J. Zimmerman, Social learning theory: A contextualist account of cognitive functioning, *Recent Advances in Cognitive-Developmental Theory* (Springer, New York, NY, 1983), pp. 1–50.
- [42] B. J. Zimmerman, Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects, *Am. Educ. Res. J.* **45**, 166 (2008).
- [43] P. D. Shen, T. H. Lee, and C. W. Tsai, Applying web-enabled problem-based learning and self-regulated learning to enhance computing skills of Taiwan’s vocational students: A quasi-experimental study of a short-term module, *Electron. J. e-Learn* **5**, 147 (2007).
- [44] J. D. Young, The effect of self-regulated learning strategies on performance in learner controlled computer-based instruction, *Educ. Technol. Res. Dev.* **44**, 17 (1996).
- [45] D. Schunk and P. Ertmer, Self-regulation and academic learning: Self-efficacy enhancing interventions, in *Handbook of Self-Regulation*, edited by M. Boekaerts, P. Pintrich, and M. Zeidner (Academic Press, San Diego, CA, 2000), pp. 631–649.
- [46] D. Kostons, T. Van Gog, and F. Paas, Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning, *Learn. Instr.* **22**, 121 (2012).
- [47] M. Komarraju and D. Nadler, Self-efficacy and academic achievement: Why do implicit beliefs, goals, and effort regulation matter?, *Learn. Individ. Differ.* **25**, 67 (2013).
- [48] S. Sungur, Modeling the relationships among students’ motivational beliefs, metacognitive strategy use, and effort regulation, *Scand. J. Educ. Res.* **51**, 315 (2007).
- [49] T. Bidjerano and D. Y. Dai, The relationship between the big-five model of personality and self-regulated learning strategies, *Learn. Individ. Differ.* **17**, 69 (2007).
- [50] A. Kitsantas, A. Winsler, and F. Huie, Self-regulation and ability predictors of academic success during college: A predictive validity study, *J. Adv. Acad.* **20**, 42 (2008).
- [51] B. K. Britton and A. Tesser, Effects of time-management practices on college grades, *J. Educ. Psychol.* **83**, 405 (1991).
- [52] S. Hurtado, D. F. Carter, and A. Spuler, Latino student transition to college: Assessing difficulties and factors in successful college adjustment, *Res. High. Educ.* **37**, 135 (1996).
- [53] H. Bembenuddy, Academic delay of gratification, self-efficacy, and time management among academically unprepared college students, *Psychol. Rep.* **104**, 612 (2009).

- [54] E. S. Indreica, A. M. Cazan, and C. Truta, Effects of learning styles and time management on academic achievement, *Proc. Soc. Behav. Sci.* **30**, 1096 (2011).
- [55] M. Richardson, C. Abraham, and R. Bond, Psychological correlates of university students' academic performance: a systematic review and meta-analysis, *Psychol. Bull.* **138**, 353 (2012).
- [56] A. Kitsantas, Test preparation and performance: A self-regulatory analysis, *J. Exp. Educ.* **70**, 101 (2002).
- [57] B. J. Zimmerman and D. H. Schunk, *Handbook of Self-Regulation of Learning and Performance* (Routledge, Taylor & Francis Group, New York and London, 2011).
- [58] B. J. Zimmermann, in *Attaining self-regulation: A social cognitive perspective*, edited by M. Boekaerts, P. Pintrich, and M. Zeidner, *Handbook of Self-Regulation* (Academic Press, San Diego, CA, 2000), pp. 13–39.
- [59] A. Bandura, *Social Foundations of Thought and Action: A Social Cognitive Theory* (Prentice-Hall, Upper Saddle River, NJ, 1986).
- [60] J. Stewart, W. Oliver III, and G. Stewart, Revitalizing an undergraduate physics program: A case study of the University of Arkansas, *Am. J. Phys.* **81**, 943 (2013).
- [61] J. P. Robinson, The validity and reliability of diaries versus alternate time use measures, in *Time, Goods, and Well-Being*, edited by J. F. Thomas and F. P. Stafford (Institute of Social Research, University of Michigan, Ann Arbor, MI, 1985), pp. 33–62.
- [62] S. A. Nonis, M. J. Philhours, and G. I. Hudson, Where does the time go? A diary approach to business and marketing students' time use, *J. Market. Educ.* **28**, 121 (2006).
- [63] R. Stinebrickner and T. R. Stinebrickner, Time-use and college outcomes, *Journal of econometrics* **121**, 243 (2004).
- [64] S. Brint and A. M. Cantwell, Undergraduate time use and academic outcomes: Results from the University of California Undergraduate Experience Survey 2006, *Teachers College record* **112**, 2441 (2010).
- [65] G. D. Kuh, *The National Survey of Student Engagement: Conceptual framework and overview of psychometric properties* (Indiana University, Center for Postsecondary Research, Bloomington, IN, 2001), accessed May 11, 2011.
- [66] J. Laing, R. Sawyer, and J. Nobel, Accuracy of self-reported activities and accomplishments of college-bound students, *J. Coll. Student Dev.* **29**, 362 (1988).