

Using time-on-task measurements to understand student performance in a physics class: A four-year study

John Stewart,^{*} Gay Stewart,[†] and Jennifer Taylor

Physics Department, University of Arkansas, Fayetteville Arkansas, 72701, USA

(Received 6 December 2011; published 4 April 2012)

Student use of out-of-class time was measured for four years in the introductory second-semester calculus-based physics course at the University of Arkansas. Two versions of the course were presented during the time of the measurement. In both versions, the total out-of-class time a student invested in the course explained less than 1% of the variance in test average and in normalized gain on the Conceptual Survey in Electricity and Magnetism. The details of how students used their out-of-class time explained from 21% to 36% of the variance in test average and 19% to 37% of the variance in normalized gain. The amount of time spent on the course increased as the course was made more difficult, but less than would be expected based on the changes in the course. Students changed their allocation of out-of-class time based on their performance within the course as the course progressed.

DOI: 10.1103/PhysRevSTPER.8.010114

PACS numbers: 01.40.Di, 01.40.Fk, 01.50.Kw

I. INTRODUCTION

Studies of time usage and behavior have been performed in many different topical fields and have a long history in more general education research. Chickering and Gamson listed time-on-task (TOT) as one of their seven principles for good practice in undergraduate education [1]. In his summary of the 320 institution, 220 000 student, National Survey of Student Engagement (NSSE), Kuh stated, “Without knowing how students spend their time, it’s impossible to link student learning outcomes to educational activities and processes associated with them” [2]. He went on to lament that of the students surveyed, 56% spent less than 15 hours per week preparing for class and 10% spent less than 5 hours. In the 2010 NSSE, the percentage of students working less the 15 hours per week had risen to 60% [3]. Time spent preparing for class is sufficiently important that it is part of the NSSE’s Level of Academic Challenge benchmark statistic.

Di Stefano [4] investigated how students use time in physics courses as part of the Introductory University Physics Project [5]. Using journals to collect data on how students use time outside the classroom, she found that on average students spent 6 hours per week on homework and that student out-of-class activities were diverse.

Welch and Bridgham studied the effect of the in-class time spent on a particular unit of study in a physics class on student achievement and found a correlation of -0.08 . They also investigated the relation between the time spent

covering a unit of instruction and the physics preparation of the class and found a correlation of -0.12 . Neither correlation was significant at the $p < 0.05$ level [6].

In a 39-study meta-analysis of the use of small group learning techniques in science, mathematics, engineering, and technology (STEM) education, Springer *et al.* found that the time students spent in the small group produced an effect size ranging from 0.52 to 0.63 on the difference in achievement, and an effect size ranging from 0.31 to 0.64 on the difference in student attitude [7].

Outside of STEM education, TOT measurements have long been used to investigate educational systems, at first traditional systems, then as computer technologies developed, to investigate computer-aided instruction, and the role of the internet in instruction. The literature on TOT is extensive; the review that follows is not meant to be exhaustive, but to demonstrate the breadth of research into student time use and the diversity of the results.

Fredrick and Walberg provided an extensive summary of early studies of the effect of time on instructional outcomes [8]. They summarize both the effect of the total time (in class and out of class) spent on instruction finding correlations ranging from 0.13 to 0.59 with performance and the details of how time was spent during the instructional day with correlations from 0.15 to 0.53. Spanjers *et al.* investigated the effect of TOT on student engagement and found that while statistically time was a significant variable, it had a weak effect on engagement [9].

In recent years, TOT has been an important part of the investigation of computer-based and internet-based learning systems. In an early 199-study meta-analysis of computer-based instruction Kulik and Kulik found a substantial decrease in instructional time resulted from the use of computer-based interventions [10]. Lou *et al.* investigated the combination of small group learning and the use of computer technology in a 122-study meta-analysis [11]

^{*}johns@uark.edu

[†]gstewart@uark.edu

and found an effect size of -0.16 for the time spent completing the small group tasks on a number of individual and group performance measures. Chen *et al.* investigated the effect of technology use on the NSSE's benchmarks for student engagement and found that the use of learning technology accounted for 0.144 of the variance in NSSE's academic challenge benchmark, of which the total time spent preparing for classes is a component [12]. Bates and Khasawneh investigated time spent online as one of the factors influencing self-efficacy in the context of web-based learning finding a negative correlation of -0.12 , so more time spent online was correlated with lower student beliefs in their mastery of the material [13].

Beyond the role of time on task, the efficacy of numerous student behaviors has been extensively studied. While reporting no summary statistics, Goldstein provides an overview of early work on the efficacy of homework in K-12 education [14]. Paschal *et al.* found an overall effect size of 0.36 in a meta-analysis of the effect of the assignment and grading of homework on performance [15].

TOT is an important control variable in investigating the efficacy of educational reform [16]. A reform activity that requires more time to complete than the instructional component it replaces may provide superior results simply because the students spend more time with the material. The U.S. Department of Education's meta-analysis of online learning provides an example of the need to control for TOT [17]. Studies where traditional face-to-face instruction provided greater or equal TOT to their online counterparts had an effect size of 0.18, while the studies where online instruction used more TOT had an effect size of 0.45.

The educational tasks performed by students and the time spent performing these tasks are often an important variable in explaining performance; however, the results for the effect of TOT vary widely, as can be seen by the preceding review, and large variations exist between studies. Physics courses offer a rich venue for continuing the study of TOT because of the variety of the activities students engage in to address the courses.

This study seeks to answer four questions:

- (1) How much of the variance in performance on hourly exams and conceptual inventories in a physics course can be explained by how students use their out-of-class time?
- (2) What features of student time use are correlated with success?
- (3) How does the time students allocate to a course change with curriculum changes?
- (4) Do students modify their use of time during the semester based on their current standing in a course?

To place the answer to the first question in context, we review the results of previous studies of the variance explained by other factors influencing performance in a

science course including logical reasoning ability, mathematics ability, previous physics preparation, high school preparation, and performance in previous mathematics and science courses.

Using the Tomlinson-Keasy-Campbell test, Liberman and Hudson [18] found $R^2 = 0.24$ for a regression of a formal operational reasoning posttest score on the final exam score. Using the same instrument as a pretest, Hudson and Liberman [19] found $R^2 = 0.19$ for a regression of the formal operational reasoning pretest against the final course grade. They also investigated the effect of mathematical reasoning ability using a mathematics pretest and found $R^2 = 0.12$ in a regression against final course grade. Wollman and Lawrenz [20] also investigated the effect of mathematical reasoning ability finding $R^2 = 0.21$ for a regression of mathematics pretest score on test average. Halloun and Hestenes [21] measured $R^2 = 0.26$ in a regression of mathematics pretest score on final course grade. They also measured $R^2 = 0.3$ for a regression of physics pretest score on course performance. Meltzer [22] found R^2 values ranging from 0.01 to 0.21 when examining the relation between mathematics pretest score and normalized gain on conceptual inventories.

Many studies also investigate the effect of combinations of factors on student performance. A regression by Hudson and Liberman [19] using both logical reasoning pretest scores and mathematical reasoning pretest scores on final course grades yielded $R^2 = 0.27$. Wollman and Lawrenz [20] measured $R^2 = 0.57$ using total grade point average (GPA), ACT score, and a mathematics pretest score when regressed upon the test average. Halloun and Hestenes [21] found $R^2 = 0.4$ using mathematics and physics pretests, $R^2 = 0.15$ using previous physics and mathematics course performance, and $R^2 = 0.49$ using mathematics and physics pretests and previous physics and mathematics course performance for regression upon final course grade. Champagne *et al.* [23] measured $R^2 = 0.325$ in a regression of logical reasoning ability, mathematical reasoning ability, and preexisting conceptual physics knowledge on a statistic formed of the average of three hourly exams and the mechanics part of the final exam. Sadler and Tai [24] investigated the effect of high school GPA, high school course decisions including the decision to take advanced placement (AP) or honors courses, and high school teaching pedagogy, among other factors. They found $R^2 = 0.26$ for regression of these factors on introductory physics grades for students who took physics in high school and $R^2 = 0.36$ for students who did not.

To summarize, mathematical reasoning ability accounts for 12%–26% of the variance in student performance in a physics class, logical reasoning 19%–24%, and physics pretest score up to 30%. Combining multiple factors can explain up to 57% of the variance in student performance.

II. RESEARCH METHODS

A. Context for research

TOT data as well as standard measures of class performance were collected for eight semesters from fall 2002 to spring 2006 in the second-semester calculus-based electricity and magnetism course at the University of Arkansas, University Physics II (UPII). This course was initially converted from traditional format to inquiry based in 1995. Its success was one of the reasons Arkansas was selected as a primary PhysTEC (Physics Teacher Education Coalition) site in 2001 and the course was used as a model for the revision of other courses as part of the PhysTEC project. The TOT data were collected as part of the monitoring of the evolution of the course. The segment of the data analyzed in this paper represents a period where the course underwent a substantial revision after the first four semesters.

The course studied uses two 50-minute lectures each week and two two-hour laboratories. The lecture uses a lecture quiz to manage attendance. The same lead instructor delivered the lecture through the period under study. Each two-hour laboratory is a mixture of demonstrations and small group work. The group work includes inquiry-based exploration, problem-solving exercises, and experiments. The course reading materials and laboratory manual were specifically constructed to support the course. Four hourly exams are given per semester. Conceptual learning is evaluated using the Conceptual Survey of Electricity and Magnetism (CSEM) as a pre and post test [25].

B. Survey design and validation

The amount of time the students were spending on different out-of-class activities was measured with two surveys that were applied after the first and third test. These surveys were constructed over the course of two semesters by a careful analysis of class policy, detailed interviews with students, analysis of students' time-use journals, and analysis of a preliminary open-ended survey instrument. As part of the survey construction process, 60 students were asked to keep time-use journals during the week preceding an hourly exam. Some activities could be deduced from course policy and our experiences with taking a science class: homework, reading, note writing, working practice tests, and attending office hours. Some activities emerged from analysis of the entries in student journals and the interviews. A preliminary survey instrument was tested during the journaling experiment. It was also found that excellent quantitative agreement between the journals and the surveys was obtained if the surveys asked the students specific, detailed questions about their time use. While no formal validity testing was performed, the experience with the journal experiment led to the construction of a survey instrument that contained very straightforward questions about actions performed and the

time required for these actions; the simplicity of the survey questions should make misinterpretation of the questions rare. Examples of the various types of survey questions may be found in Appendix B.

Survey development was informed by previous work examining the validity of surveys asking respondents to self-report factual data. The factual information used in this study included whether an action was performed, how many times it was performed, and how much time it took. Laing *et al.* found only a 10% error rate in self-reported data asking for factual responses [26].

Kuh summarized the characteristics of survey questions which produced accurate self-reported responses [27]: the information must be known to the student, the questions must be clearly worded, the information requested must be recent, students must take the questions seriously, and answering should not embarrass the students. To meet these criteria, the surveys asked questions about tangible actions taken by the students, information all the students should possess. The surveys were administered as soon as possible in laboratory during the week after the exam. Questions were simply worded and the students were told the information was appreciated for planning the course in future semesters. While surveys could not be completely anonymous because results had to be associated with test and conceptual inventory performance, the students were told that no one associated with the course would see data that identified the student and that nothing would be done with the surveys until final grades for the course were entered.

C. Course revision

In the period of time analyzed in this paper, two versions of the course were offered. The original version, version I, was offered from fall 2002 to spring 2004, and version II from fall 2004 to spring 2006. The course was substantially revised for the fall 2004 semester. Additional conceptual material was added to fully coordinate the lecture or reading segment of the course with the laboratory segment. Additional material emphasizing the use of calculus was added both to the reading and to problem-solving activities in laboratory. These changes increased the number of topics covered, the amount of reading required, and the mathematical difficulty of the material covered. Students were required to master more material and to master that material with increased mathematical facility.

Conversations with students showed they found the revised course harder and the test average dropped from 80% for version I to 74% for version II. All tests in the period studied were written by the same lead instructor. He feels the test difficulty increased because of the increased number of topics covered by each test and the increase in the mathematical sophistication of the coverage of some topics. As a measure of the amount of additional material in the second version of the course, the course readings

grew from 98 000 words in version I to 140 000 words in version II. Version I produced a CSEM posttest score of 61% and a normalized gain of 0.43; version II produced a posttest score of 65% and a normalized gain of 0.49. The lead instructor for the period studied attributes these increases to increased conceptual coverage in the course reading material and the homework and the redesign of some inquiry-based and small group activities in the laboratory portion of the course.

D. Data collection

Once constructed, the survey instruments were applied to students in UPII. The surveys have been administered continuously since fall 2002, but this study analyzes the data from the fall 2002 semester to the spring 2006 semester.

In version I of the course, a total of 455 students completed the course for the four semesters studied. Of the 455 students, 377 completed both the CSEM pre and post test allowing a normalized gain to be calculated. Of the students completing the course for which a normalized gain could be calculated, 194 completed both surveys with all questions answered. These students will be included in the analysis which follows. In version II, a total of 364 students completed the course, a normalized gain could be calculated for 313 students, and 166 of these students completed both surveys answering all questions.

The students included in the study are a somewhat different population than the class as a whole. The students included had to be in attendance in the laboratory sessions where the pretest and posttest were given and in the laboratory sessions where the two surveys were given, so students included in the analysis may be expected to on average have higher attendance than the average student in the class. The difference in the two populations will be investigated in Sec. III A.

E. Survey questions

The two surveys contained a variety of questions that asked students what activities they did, how many times they did them, how much time they spent doing them, and when the activities were performed in relation to other activities. The following summarizes the two surveys and details of the coding of non-numerical responses. For yes or no questions, “no” was coded as 0 and “yes” as 1. When coding non-numeric responses, higher numbers were used for behavior that the lead instructor of the course felt was more likely to lead to superior performance.

1. Survey 1

Survey 1 was given the week following the first hourly exam. It asked how many times the assigned readings were read before the test, when the material was first read (before lecture = 4, while working the reading self-test = 3,

while working the homework = 2, before the exam = 1, never = 0), the time per homework set, and the study time for exam one. Reading was assigned to be completed before the lecture in which the material was covered. To encourage completion of the reading assignment, each homework contained a set of questions on the assigned reading, the reading self-test.

Three practice tests are published for each exam. Students were asked whether they read, worked, or read the solution for each of the three tests. The students were asked if they took lecture notes and, if so, if the notes were used to prepare for the exam. Students were also asked how many of the homework solutions were reviewed while studying for the exam.

2. Survey 2

Survey 2 requested a more detailed accounting of time use and was given the week after the third hourly exam. Students were asked how often they took lecture notes (always = 3, sometimes = 2, rarely = 1, never = 0). If lecture notes were taken, were they used while working homework and while studying for the exam. The students were asked how many of the homework solutions were reviewed while studying for the test. The students were asked when they first did the course reading (the same coding was used as in survey 1) as well as how many times the required reading was done before the exam. They were asked for the amount of time spent reading excluding studying for the exam, the average time spent working a homework set, and their total study time for the last exam. Survey 2 collected the same information as survey 1 on the practice tests. The students were then asked to account for how their study time was divided among reading the required text, working examples in the reading, reading the study guide, reading other sources, working practice tests, reviewing homework solutions, reviewing lecture notes, reworking homework problems, asking questions or getting help, and other. Finally, the students were asked again how much time they spent preparing for the exam.

We note that the students were asked for their total study time twice on survey 2, once on the front page, and then again on the back page after being asked to itemize their time use. The reported times for the two questions were quite consistent with an average difference per student of 0.75 hours for version I and 0.5 hours for version II.

Examples of the survey questions can be found in Appendix B.

F. Composite variables

Some additional variables were calculated from the responses to the two surveys. All responses involving the use of the practice tests from both surveys were combined into a single variable, “practice test usage,” where the total number of actions performed was divided by the actions possible. The total study time, “study time,” was

calculated by averaging the response on survey 1 and the two responses on survey 2. The homework time per week, “homework time,” was calculated by averaging the time per set from the two surveys, then computing the total time per semester using the number of homework sets turned in for each student, and then dividing by the number of weeks in a semester. The time spent reading in a nonexam week, “reading time,” was taken from survey 2. An estimate of time use during a nonexam week, “normal weekly time,” was formed by adding the homework time to reading time. An average out-of-class time per week, “total time,” was found by adding one-quarter of the study time to the normal weekly time. The most important composite variables are summarized in Table I.

G. Outliers

Some survey questions asked the student to report a number as the total time to perform some activity. This presented a temptation to some students to greatly over-report their time use. These responses would have undue influence on the statistical analysis. To eliminate these inflated responses, Gaussian distributions were fit to the responses to the open-ended questions. Students with responses outside the fitted curve were eliminated from the study. Such students’ responses were more than 3 standard deviations from the mean. This resulted in the removal of 8 students from version I and 19 students from version II. With these modifications, for version I, the maximum reported total time was 15 hours and the maximum study time was 21 hours. For version II, the maximum total time was 19 hours and the maximum study time was 25 hours. While large for a four credit course, these times are not inconsistent with those reported to the lead instructor by students who are struggling with the material.

H. Statistical analysis

All statistical results presented were produced using the SAS statistics software. Simple and multivariate linear regression analysis were performed using the REG procedure and correlation analysis using the CORR procedure. To aid comparison with results summarized in the Introduction, both the correlation coefficient r and the amount of variance explained by the linear model R^2 are reported. These statistical techniques were selected because they allowed the most natural comparisons with the other studies of factors affecting student performance

detailed in the Introduction. When correlation analysis was used, R^2 is calculated as r^2 . The statistics used are discussed in more detail in Appendix A.

III. RESULTS

A. Effect of missing assignments

One feature of student behavior can be examined without the use of the two surveys, the effect of missing an assignment. The total number of assignments not turned in, the total missed assignments, was calculated from the record of student grades by summing the number of homework assignments, lab quizzes (laboratory quizzes), and lecture quizzes missed. The “Total” rows of Table II present the correlation coefficient, r , and R^2 resulting from correlation analysis relating the total missed assignments with test average and normalized gain. The “All” columns report results for all students who completed the course, $N = 455$ for version I and $N = 364$ for version II for test average and $N = 377$ and $N = 313$ for normalized gain. The “In study” columns present the same statistics only for students who completed the course, the pre and post test, and answered all questions on the two surveys; those students who are included in the analysis of later sections.

For version I of the course, the test average of the full class was 80%, while the test average of the students retained for this study was 82%; the normalized gain for the full class was 0.43 and 0.45 for the students in this study. For version II of the course, the test average of the full class was 74%, while the test average of the students retained for this study was 77%; the normalized gain for the full class was 0.49 and 0.51 for the students in this study. Therefore, the students included in the study were slightly more successful in the class than the general population of students.

The test performance of the full set of students was substantially more dependent on the number of missed assignments than the test performance of students who will be included in further analysis. As further analysis is performed, it will be found that the $R^2 = 0.10$ and 0.14 for test average are large compared to the R^2 of other variables and the primary behavior a student could exhibit that is correlated with success is turning in the work in the class. This large effect of missed assignments supports the findings of Palazzo *et al.* that homework copying has a significant negative impact on course performance [28]. A

TABLE I. Summary of composite variables.

Homework time	The homework time per week averaged over the semester.
Reading time	The total out-of-class time spent reading for weeks not containing an exam.
Study time	The average time spent preparing for an exam.
Normal weekly time	The total out-of-class time for weeks not containing an exam.
Total time	The average time per week including the time spent preparing for exams.

TABLE II. The effect of missed assignments on performance. For version I—All, $N = 455$ students are included in correlations with test average, and $N = 377$ students are included in correlations with normalized gain. For version II—All, $N = 364$ students are included in correlations with test average, and $N = 313$ students are included in correlations with normalized gain. For version I—In study, $N = 194$ students are included in correlations with test average and normalized gain. For version II—In study, $N = 166$ students are included in correlations with test average and normalized gain.

Version I—Test average				
Missing	All		In study	
	r	R^2	r	R^2
Homework	-0.285 ^b	0.081 ^b	-0.221 ^a	0.049 ^a
Lecture	-0.231 ^b	0.053 ^b	-0.264 ^a	0.070 ^a
Lab	-0.280 ^b	0.079 ^b	-0.142 ^a	0.020 ^a
Combined		0.106 ^b		0.076 ^a
Total	-0.321 ^b	0.103 ^b	-0.269 ^b	0.073 ^b

Version I—Normalized gain				
Missing	All		In Study	
	r	R^2	r	R^2
Homework	-0.091	0.008	0.023	0.001
Lecture	-0.087	0.008	-0.015	0.000
Lab	-0.088	0.008	-0.024	0.001
Combined		0.012		0.003
Total	-0.115 ^a	0.013 ^a	-0.009	0.000

Version II—Test average				
Missing	All		In Study	
	r	R^2	r	R^2
Homework	-0.344 ^b	0.118 ^b	-0.261 ^a	0.068 ^a
Lecture	-0.322 ^b	0.104 ^b	-0.290 ^b	0.084 ^b
Lab	-0.339 ^b	0.115 ^b	-0.176 ^a	0.031 ^a
Combined		0.151 ^b		0.099 ^a
Total	-0.379 ^b	0.144 ^b	-0.302 ^b	0.091 ^b

Version II—Normalized gain				
Missing	All		In Study	
	r	R^2	r	R^2
Homework	-0.181 ^a	0.033 ^a	-0.206 ^a	0.043 ^a
Lecture	-0.260 ^b	0.068 ^b	-0.291 ^b	0.085 ^b
Lab	-0.195 ^a	0.038 ^a	-0.115	0.013
Combined		0.076 ^b		0.086 ^a
Total	-0.246 ^b	0.060 ^b	-0.263 ^a	0.069 ^a

^aSignificant at the $p = 0.05$ level.
^bSignificant at the $p = 0.0001$ level. For all entries except the “Combined” entry, R^2 is estimated as r^2 . For the “Combined” entry, multivariate linear regression was performed using three variables, the missing homework, lectures, and laboratories. The R^2 reported is the R^2 for this regression.

student who copies a homework assignment in effect misses the assignment because he or she does not invest the intellectual effort to solve it.

Table II also presents the effect of missing assignments of different types. While there is some variation by assignment type, the effect of the three types of assignment is fairly consistent. The “Combined” row reports the R^2

resulting for a linear regression using the number of missed assignments of each type as independent variables. The combination of the three variables often explained much more variance than any individual variable, but in most cases the R^2 for the three variables was quite similar to R^2 for the total number of missing assignments.

This negative effect of missing assignments on performance is consistent with the work summarized by Fredrick and Walberg [8] who find a correlation up to 0.59 with the time spent on instruction. A student who misses a homework lowers his or her personal instructional time. A student who misses a lab or lecture quiz has missed the class meeting in which the quiz was given. The above also supports the observation of Paschal *et al.* of a positive effect size with the assignment and grading of homework [15].

B. Effect of total time on task

The results of regression and correlation analysis of the total out-of-class time a student invested in the course is presented in Table III. Normal weekly time is the out-of-class time spent during a nonexam week, study time is the average time spent preparing for an exam, and total time is the average time spent per week averaging the exam preparation time over the semester. The Total missed assignments is included to provide a reference for the size of the correlations.

The total time is not significantly correlated with either test average or normalized gain. In fact, the correlation coefficients for total time are all negative, indicating that students who spend more time on the course perform more weakly. Three of the four correlations of normal weekly time were positive, indicating more time spent in weeks not including an exam led to better performance, but none of the correlations were significant. Three of the four correlations for study time were significant at the $p = 0.05$ level, but all correlation coefficients for this variable were negative. The more time a student spent preparing for tests the worse he or she performed on both the tests and the CSEM.

At first, these results are quite counterintuitive; working harder in a course or studying harder for an exam should lead to improved results. These results can be understood by remembering that a college student is not generally trying to produce the best possible performance in a single class, but is trying to achieve a certain set of goals in a number of classes along with maintaining enough time to participate in a number of nonclass activities. A student, then, may choose to invest only the time needed to achieve his or her performance goal which would mean particularly capable students would choose to use their facility with the material to lower the time spent on the course and students struggling with the course would increase their time commitment.

This effect of spending more time on intellectually valuable tasks resulting in lower performance is seen in a

TABLE III. Time on task. The variables r_{test} and R^2_{test} measure the correlation and R^2 with respect to test average; the variables r_{gain} and R^2_{gain} measure the correlation and R^2 with respect to the normalized gain. For version I, $N = 194$ students were included, and version II included $N = 166$.

Variable	Version I				Version II			
	r_{test}	R^2_{test}	r_{gain}	R^2_{gain}	r_{test}	R^2_{test}	r_{gain}	R^2_{gain}
Total missed assignments	-0.269 ^b	0.073 ^b	-0.009	0.000	-0.302 ^b	0.091 ^b	-0.263 ^a	0.069 ^a
Normal weekly time	0.041	0.002	0.022	0.001	-0.020	0.000	0.048	0.002
Study time	-0.150 ^a	0.022 ^a	-0.244 ^a	0.060 ^a	-0.121	0.015	-0.260 ^a	0.068 ^a
Total time	-0.024	0.001	-0.075	0.006	-0.071	0.005	-0.084	0.007

^aSignificant at the $p = 0.05$ level.
^bSignificant at the $p = 0.0001$ level.

number of the studies presented in the Introduction. Welch and Bridgham saw a negative correlation between the time spent on a unit of study and performance [6]. Bates and Khasawneh also measured a negative correlation between the time spent using web-based learning sites and self-efficacy [13].

To investigate the effect of how the allocation of time among a number of primary course activities affects performance, the ratio of the time spent performing common activities with total time and study time were calculated and regression and correlation analysis repeated. The results of this analysis are shown in Table IV. Reading time is the average time spent reading in a nonexam week. Practice test time is the amount of study time spent reviewing the practice tests. Study time—reading is the amount of study time spent reading. Maximum study item is the maximum amount of study time spent performing one of the itemized study actions in survey 2. The ratio of maximum study item to study time measures the degree to which a student’s exam preparation is focused into a single activity.

The pattern of results found in Table III was repeated with only the ratio of study time to total time having significant correlations with test performance or normalized gain for three of the four cases. Once again, all correlations of this variable are negative so the greater fraction of a student’s time that is spent preparing for the exam the weaker both the test performance and the conceptual gains. The ratio of maximum study item to study

time was significantly positively correlated with test average and normalized gain in the harder version of the course indicating that focusing more exam preparation into a single study activity was a winning strategy for the harder course. The ratio of the amount of study time spent reading to the total study time was significantly positively correlated with normalized gain for both versions of the course.

C. Correlation of student activities with performance

The average time spent on the course explained little of the variation in performance on either exams or the CSEM. The two surveys collected detailed information on both the activities performed for the course and the time spent in performance of those activities. Table V presents the five most highly correlated variables for both versions of the course.

For both versions of the course, the total missed assignments was the most strongly correlated variable with test performance. Lecture attendance, measured by the fraction of lecture quizzes turned in, and the homework completion rate, the fraction of homework sets turned in, two variables that contribute to total missed assignments, were also among the top five most strongly correlated with test average.

The remaining of the five most strongly correlated variables with test average were two of the itemized study activities, getting help and working examples in the course reading, and the ratio of study time to total time. The correlations of all three variables with test average were

TABLE IV. Division of time on task among activities. For version I, $N = 194$ students were included, and version II included $N = 166$.

Variable	Version I				Version II			
	r_{test}	R^2_{test}	r_{gain}	R^2_{gain}	r_{test}	R^2_{test}	r_{gain}	R^2_{gain}
Study time/total time	-0.202 ^a	0.041 ^a	-0.224 ^a	0.050 ^a	-0.119	0.014	-0.280 ^a	0.079 ^a
Reading time/total time	0.046	0.002	0.161 ^a	0.026 ^a	-0.031	0.001	0.141	0.020
Maximum study item/study time	0.132	0.017	0.120	0.014	0.161 ^a	0.026 ^a	0.207 ^a	0.043 ^a
Practice test time/study time	0.052	0.003	-0.005	0.000	0.097	0.009	0.243 ^a	0.059 ^a
Study time—Reading/study time	0.019	0.000	0.170 ^a	0.029 ^a	0.108	0.012	0.167 ^a	0.028 ^a

^aSignificant at the $p = 0.05$ level.

TABLE V. Five most highly correlated activities.

Version I—Test average		
Activity	r	R^2
Total missed assignments	-0.269 ^b	0.073 ^b
Lecture attendance	0.267 ^a	0.071 ^a
Homework completion rate	0.226 ^a	0.051 ^a
Study time/total time	-0.202 ^a	0.041 ^a
Study time—working	-0.164 ^a	0.027 ^a
Chapter examples		
Version I—Normalized gain		
Study time	-0.244 ^a	0.060 ^a
Study time—review homework	-0.234 ^a	0.055 ^a
Study time/total time	-0.224 ^a	0.050 ^a
Study time—practice test	-0.222 ^a	0.049 ^a
Study time—review lecture notes	-0.181 ^a	0.033 ^a
Version II—Test average		
Total missed assignments	-0.302 ^b	0.091 ^b
Lecture attendance	0.290 ^b	0.084 ^b
Homework completion rate	0.261 ^a	0.068 ^a
Study time—working chapter examples	-0.249 ^a	0.062 ^a
Study time—getting help	-0.205 ^a	0.042 ^a
Version II—Normalized gain		
Homework solutions reviewed	-0.356 ^b	0.127 ^b
Study time—review homework	-0.311 ^b	0.097 ^b
Lecture attendance	0.291 ^b	0.085 ^b
Study time/total time	-0.280 ^a	0.079 ^a
Total missed assignments	-0.263 ^a	0.069 ^a

^aSignificant at the $p = 0.05$ level.

^bSignificant at the $p = 0.0001$ level.

negative. Again, students who invest more time of their total time in exam preparation are less successful on the tests. The negative correlations with getting help and working examples may indicate these behaviors are more commonly performed by students who are still struggling to understand the material while they study, as opposed to students who have already developed an understanding of the material by the time they begin to prepare for the exams and spend their exam preparation in practicing and refining their understanding.

The variables most strongly correlated with test average were not generally among the variables most strongly correlated with normalized gain. There was little consistency in the five variables most strongly correlated with the normalized gain. The review of homework solutions during exam preparation measured both in the time spent and in the number of solutions reviewed was negatively correlated with normalized gain. The ratio of study time to total time, again negatively correlated, was also one of the five variables most correlated with normalized gain in both versions.

The primary result that emerges from this analysis is that no behavior explained more than 7% of the variance in test average or 12% of the variance in normalized gain; there is no particular behavior that guarantees success in a physics

course. In general, students who turn in their assignments, attend lecture, and spend a larger fraction of their total class time preparing during nonexam weeks did better in the course.

D. Variance in performance explained by behavior

The previous sections investigated the effect of individual behaviors, but how much of the variation in student performance can be explained by how a student combines various behaviors? To investigate the total variance explained by behavior, a multivariate linear regression analysis was performed and is presented in Table VI. The model selected was the model that maximized R^2 adjusted, R^2_{adj} , a statistic that corrects the variance explained for the loss of degrees of freedom associated with adding variables to the linear model. In all cases the maximum R^2_{adj} model required many variables with the last variables added to the model explaining little additional variance. As such, Table VI presents the model that explains 50%, 75%, and 100% of the variance explained by the maximal R^2_{adj} model. The variables used in the 50% and 75% models are reported. Variables marked with an “*” had negative coefficients.

The maximal R^2_{adj} models explained 20.5% and 35.7% of the variance in test average in versions I and II and 18.7% and 37.0% of the variance in normalized gain. All of the 50%, 75%, and 100% models were significant at the $p < 0.0001$ level except the 50% model for the version I normalized gain where $p = 0.0002$.

At 75% of the maximal R^2 , 4 variables explained 15% of the variance in test average in version I and 6 variables 25% of the variance in version II. The study behavior of working examples in the text had a negative coefficient in both versions as did the total number of missed assignments. No other variables occurred in the linear models for both versions for the test average.

At the 75% level, 3 variables explained 11% of the variance in normalized gain in version I and 4 variables 27% of the variance in version II. The variables used in these models were quite different than the variables used in the test average models. The only variable in common for both versions was the fraction of the exam preparation time that was spent reading; this had a positive coefficient in both models.

Either using the variance explained by the maximal R^2_{adj} models or the 75% models, the amount of variance explained by how students manage their time had similar explanatory power to features such as logical reasoning ability or mathematical reasoning ability examined in the Introduction.

For both test average and normalized gain, the details of student behavior explained substantially more variance in the second version of the course. In the harder course, how students invest their time becomes more important.

TABLE VI. Variance in test average and normalized gain explained by behavior. For each version and for the test average and normalized gain, the last model reported is the model that maximizes R^2_{adj} . The previous two models have R^2 that is 50% and 75% of this maximal model. Entries marked with “*” have negative coefficients.

R^2 %	Parameters	R^2	R^2_{adj}	Model
Version I—Test Average				
50%	2	0.103 ^b	0.094	*Study time—review homework, *Total missing assignments
75%	4	0.153 ^b	0.135	*Study time—work example, *Total missing assignments, *Total time, Total reading time
100%	10	0.205 ^b	0.162	
Version I—Normalized gain				
50%	2	0.086 ^a	0.077	*Study time, Study time—reading/study time
75%	3	0.107 ^b	0.093	*Study time—review homework, *Study time/total time, Study time—reading/study time
100%	10	0.187 ^b	0.143	
Version II—Test average				
50%	3	0.186 ^b	0.171	*Study time—work example, When first reading completed, *Total missing assignments
75%	6	0.251 ^b	0.222	*Study time—work example, *Study time—review homework, When first reading completed, *Use lecture notes while studying, *Study time—get help, *Total missing assignments
100%	15	0.357 ^b	0.293	
Version II—Normalized gain				
50%	2	0.197 ^b	0.187	*Solutions reviewed for exam, Lecture attendance
75%	4	0.270 ^b	0.252	*Solutions reviewed for exam, Lecture attendance, Study time—practice test/study time, Study time—reading/study time
100%	17	0.370 ^b	0.297	

^aSignificant at the $p = 0.05$ level.
^bSignificant at the $p = 0.0001$ level.

E. Reaction of student time use to curricular modifications

The curricular change between version I and version II increased the amount of assigned reading and the mathematical difficulty of the course and the homework. While the number of homework problems per week was approximately constant, the difficulty and coverage of the homework increased between version I and version II. Table VII summarizes the total time use for the two versions.

Table VII does show an overall increase in total time from version I to version II, but of only about 1/2 an hour per week. This additional time was the result of increased

time preparing for exams. The amount of time invested in a week not containing an exam was unchanged. The students did spend slightly more time reading, but actually less time working on homework. This suggests that the amount of extra time an instructor can extract from a student without giving additional graded assignments is fairly limited. The more difficult version II pushed the students into behavior that the correlation analysis suggests was counterproductive in that the students spent a larger fraction of their out-of-class time in exam preparation.

F. Modification of behavior during semester

The previous section examined modifications to the total time committed to the course due to a curricular change. The time students spent in weeks not containing an exam was unchanged between the two versions of the course suggesting that the amount of time committed to a course may be relatively inflexible, depending on factors outside the course design. This section investigates whether this inflexibility extends to student time use during a course. Table VIII shows the average amount of time invested in preparing for the first hourly exam and the change in the amount of time invested between the first and third exams. These averages are separated by the student’s test average at the time of the third exam. The first exam covers

TABLE VII. Average out of class time use.

Activity	Version I (hours)	Version II (hours)
Normal weekly time	4.43	4.43
Normal weekly reading time	1.66	1.86
Normal weekly homework time	2.77	2.57
Study time	5.71	7.35
Study time—reading	1.06	1.05
Total reading time	1.96	2.12
Total time	6.06	6.53

TABLE VIII. Change in study behavior during semester. Shown are the number of hours spent preparing for hourly exams separated by the test grade of the student going into the third exam. Also shown is the difference in the number of hours spent preparing for exam 3 and exam 1. The number in parenthesis is the number of students included in the average.

	Exam 1 time (hours)			
	A	B	C	D-F
Version I	4.0 (85)	5.0 (76)	6.4 (21)	3.7 (12)
Version II	4.1 (48)	5.7 (66)	6.5 (33)	6.5 (19)
	Exam 3 time—Exam 1 time (hours)			
	A	B	C	D-F
Version I	1.2 (85)	1.2 (76)	1.1 (21)	3.4 (12)
Version II	1.8 (48)	3.0 (66)	3.6 (33)	1.5 (19)

somewhat less material than the third exam with one-half of a week less lecture. The material on the third exam, magnetism, is more mathematically challenging for the students than the material on the first exam, electrostatics, so some increase in exam preparation time should be expected.

The pattern of changes in student time commitment from exam 1 to exam 3 was quite different for the two versions of the course. For version I, the time spent studying for the first exam consistently increased as the test grade on the first two exams decreased, except that the students earning a “D” or “F” on the first two exams invested less time than other students. The “A” to “C” students increased their preparation time by about the same amount of time for exam 3, with the “D”-“F” students increasing the study time much more than other students. The third test is given after the drop deadline, so students who are failing the class but decide to remain in the class should and did greatly increase their time commitment to the class. Version II also showed an increase in study time for exam 1 as grades decreased, but did not have the dip in study time for the lowest performing students. The amount of additional time invested for exam 3 was much higher and also increased with decreasing test grade; however, the lowest performing students actually increased their time investment less than other students.

The observed pattern could be explained if students in the easier version of the course were receiving the grade they desired, so the increase in effort tracked the increase in difficulty from exam 1 to exam 3. In the harder version of the course, some students may have had a lower grade than desired and increased their effort to reach their goal.

It seems there was a subpopulation of students in the first version of the course who were not investing a serious effort in the course. All students seemed to be investing significant time in the second version of the course, but the lowest performing students did not invest additional effort commensurate with the rest of the class. One possible explanation of this observation is that for the harder course

the weaker students might have been resigned to doing badly in the course.

IV. DISCUSSION

Four research questions were proposed in the introduction; we will examine them in the order proposed.

(1) *How much of the variance in performance on hourly exams and conceptual inventories in a physics course can be explained by how students use their out-of-class time?* The total out-of-class time (total time) students invested in the course explained little of the variance ($R^2 < 0.01$) in either test average or normalized gain on the CSEM.

Of the variables measuring overall time use, the time spent preparing for exams was most consistently significantly correlated with performance. This variable explained 6%–7% of the variance in normalized gain, but only 1%–2% of the variance in test average. In our study, the overall commitment of out-of-class time is of less importance in explaining performance than the factors detailed in the Introduction: logical reasoning ability, mathematics reasoning ability, or physics preparation.

To examine how the relative division of time use among activities affected performance, ratios of the most general time-use variables were investigated. The ratio of study time to total time was the most productive in explaining variations in performance. It explained 1%–4% of the variation in test average and 5%–8% of the variation in normalized gain, also smaller than the factors summarized in the Introduction.

Multivariate linear regression analysis allowed the investigation of the effect of multiple variables. A combination of variables explained much more of the variation in student performance, explaining 21%–36% of the variance in test average and 19%–37% of the variance in normalized gain. This level of explanatory power is comparable or superior to the variance explained by mathematical reasoning, logical reasoning, or physics preparation in previous studies.

The maximal R_{adj}^2 models used 10–17 time variables and were slowly converging. As such, linear models explaining a fixed fraction of the variance explained by the maximal R_{adj}^2 model were investigated. For models with R^2 of 75% of the maximal R_{adj}^2 model, 4–6 variables explained 15%–25% of the variance in test average and 3–4 variables explained 11%–27% of the variance in normalized gain. The amount of variance explained by behavior was much higher in the harder version II course. The 11%–27% of the variance explained by the 75% models places student behavior as a factor of commensurate importance to mathematical or logical reasoning ability, but somewhat less important than physics pretest score. This is particularly impressive because the time-use data collected by the two surveys does not contain any graded measure of student ability such as a pretest score or a homework average.

(2) *What features of student time use are correlated with success?* In both versions of the course, the variable most strongly correlated with test performance was the total number of missed assignments which explained 7%–9% of the variance in test average on its own. The second and third most highly correlated variables were the lecture attendance rate and the homework completion rate which were used to calculate the total number of missing assignments. The other variables most highly correlated with test average all involved details of time use during exam preparation and were all negatively correlated suggesting working chapter examples, getting help during exam preparation, or investing a higher fraction of class time in exam preparation are not effective preparation strategies for the course under investigation.

The variables most strongly correlated with normalized gain were very different. Variables measuring attendance or assignment completion were among the most important only in version II of the course. Only the amount of study time spent reviewing homework solutions and the ratio of study time to total time were among the top five most strongly correlated variables in both versions of the course.

Examination of the 75% linear regression models provides additional insight. For test average, the total number of missed assignments appears as a variable selected in both versions of the course. Also represented in these models are a variety of exam preparation behaviors with negative coefficients, working chapter examples, reviewing homework, and getting help. The variables with positive coefficients are the relative timing of the first reading of the course materials and the total time spent reading. The positive coefficient of the placement of the first reading indicates that students who do their reading earlier perform better on the test.

The picture is quite different for normalized gain. The importance of missed assignments is substantially reduced with only lecture attendance selected as a variable for the linear model and then only in version II of the course. The fraction of exam preparation time spent reading was selected as a variable with positive coefficient for both versions of the course.

Overall, the time spent reading and when that reading occurs stands out as a positive factor influencing test performance and normalized gain.

(3) *How does the time students allocate to the course change with curriculum changes?* There was an increase in the total out-of-class time devoted to the course between version I and version II of the course; however, all the increase in time went into additional exam preparation time. The average time spent in a nonexam week was the same for both courses. While students did modify their time commitment because of the curricular reform, the additional one-half hour per week invested was far less than suggested by the extent of the reform.

(4) *Do students modify their use of time during the semester based on their current standing in a course?* Students did change the amount of time invested in exam preparation between the first and third exam. The amount of additional time invested for the more challenging third exam was dependent on the students' test average through the first two exams. The pattern of time-use changes was markedly different for the two versions of the course.

V. IMPLICATIONS FOR INSTRUCTION

Preliminary analysis of this data was presented at the 2006 summer American Association of Physics Teachers meeting. The above represents a reworking of the analysis which eliminated all students with missing information. The lead instructor of the course used the preliminary results, which are consistent with the above, to inform the advice he gives to students. The results above also have implications for overall course design.

For students, the main point emerging from this analysis is that attending class and turning in assignments are actually the most important things they can do to be successful in the course. Beyond this obvious advice, the above indicates that students who spend a larger fraction of their time in preparation for the course before they begin to study for the exams perform better. The instructor's new primary advice to students who are not performing well in the course, but are turning in assignments, is that they must invest more time earlier in nonexam preparation activities such as reading and working on homework, that they must develop their understanding of the material earlier. This is further supported by the negative correlations of exam preparation behavior indicating a student is still developing understanding rather than refining understanding while studying such as reviewing homework solutions or working examples in the text.

The results of this study are very informative for the design of a science course and for an instructor's understanding of the functioning of a course. The very weak effect of total time investment is an indication that students are managing time across their entire college experience and performance in one class cannot be viewed in isolation. Upon viewing the regression results for total time, the lead instructor of the course could produce numerous examples where good students performed more weakly than expected because they had to split their exam preparation time over multiple courses.

The consistency of the time invested in nonexam weeks between the two versions of the course seems to indicate that the amount of time available to a course is an intrinsic feature of the student population. As the course was made harder, the students did not change their behavior to match the change. Students did invest more time studying for the harder course, but as previous discussion shows this

additional effort would be more effectively directed if it were invested in a nonexam week. The harder course also seemed to leave the weakest performing students behind; whereas the weakest students seemed to coast through the first exam in the easier version and then greatly increase their effort for later exams.

Students did seem to adjust their exam preparation time to the difficulty of the material and to their current standing in the course. The amount of adjustment was dependent on the difficulty of the course. Since the students do regulate time investment interior to a course, some variation of course time requirements within the semester should be acceptable.

VI. LIMITATIONS

This study uses correlation analysis to investigate the effects of student behavior on performance. Naturally, correlation does not imply causation, so it is impossible to differentiate between behaviors that cause superior performance and behaviors that are naturally carried out by higher performing students.

The primary conclusions of the study involve the traditional behaviors of attending lecture and turning in homework. Much research has shown that interactive-inquiry methods are required in many cases for achieving conceptual understanding and overcoming misconceptions [29]. In the course under study, the inquiry-based instruction occurs solely in the laboratory component of the course. The results in Table II suggest that the laboratory component of the course is of equal or lesser efficacy to the more traditional lecture and homework components. The lower than expected correlations of the laboratory component are quite probably the result of an extremely strong laboratory attendance policy that produces very high laboratory attendance with most students in the study missing one or zero of the 28 laboratories. This means the variation in the laboratory attendance is far lower than the variation in lecture attendance or homework completion rate. As such, it is unwise to draw conclusions about the relative efficacy of traditional and reform instruction from the results of this study.

VII. FUTURE

The work presented in this study was limited to the students in one course at a single institution. The study should be extended to additional courses and institutions to see if the conclusions are consistent and to develop an understanding of the range of variation of the effect of time on performance. The data taken on the course studied from 2006 to 2011 will be used to investigate the effect on time usage of the implementation of a computer-based homework system and a computer-based self-testing tool. The study will be extended in the future to collect data on additional factors which affect student performance and

potentially impact time use such as native student ability measured by ACT score, high school physics preparation, and mathematical preparation. Ideally, one would extend this research to study some of the correlations found by dividing the class into an experimental group and a control group and adjusting the assignments of the experimental group to modify their time use.

VIII. CONCLUSION

The total time spent preparing for a physics course, the average time spent in a nonexam week, and the average time spent preparing for an exam explained 0%–2% of the variance in test average and 0%–7% of the variance in normalized gain. The number of missed assignments was consistently the most important factor influencing test average, but was less important for the normalized gain on the CSEM. Both the amount and timing of reading were positively correlated with performance in the multiple regression models. Students increased the time devoted to the course when the course was made more difficult; however, this increased effort was only evident in an increase in exam preparation time and seemed to be insufficient for the amount of additional material added to the course.

ACKNOWLEDGMENTS

This work is supported in part by the National Science Foundation as part of the evaluation of improved learning for the Physics Teacher Education Coalition, Grant No. PHY-0108787.

APPENDIX A: STATISTICS

This paper used correlation analysis, linear regression analysis, and multivariate linear regression analysis. The strength of the correlation was estimated using r , the degree to which a linear model explained the data was measured by R^2 , “ R squared,” and linear models were selected that maximized R^2_{adj} , “ R squared adjusted.” This appendix presents a brief discussion of these statistics following Devore [30].

Linear regression analysis seeks to extract a linear model from a set of data pairs (x_i, y_i) , where x_i is assumed to be an independent control variable and y_i is a dependent variable that for a given x varies around an average value \bar{y}_x . Linear regression extracts a model $y = \beta_0 + \beta_1 x$ that allows the calculation of a predicted \hat{y}_i value for each x_i . The total variation in y of the data set can be characterized using the total sum of squares (SST), where $\text{SST} = \sum_i (y_i - \bar{y})^2$, where $\bar{y} = \sum_i y_i / N$ is the average of y and N is the number of data points. The variation remaining after the assumption of a linear model, the squared vertical distances from the regression line, is captured by the error sum of squares (SSE), $\text{SSE} = \sum_i (y_i - \hat{y}_i)^2$. The coefficient of determination, R^2 , is then $R^2 = 1 - \text{SSE}/\text{SST}$, the

fraction of initial variation explained by the assumption of a linear model.

Correlation analysis does not assume an independent and dependent variable but uses (x_i, y_i) as a pair of observations where both x_i and y_i vary according to some probability distribution. The correlation coefficient for the distribution is defined as $\rho = \text{Cov}(x, y)/(\sigma_x\sigma_y)$, where σ_x is the standard deviation in x , σ_y is the standard deviation in y , and $\text{Cov}(x, y)$ is the covariance. The statistic r is an estimate of ρ from the data set. If a linear model were fit to the data set, the R^2 calculated for the linear regression would equal r^2 .

Multivariate linear regression generalizes the above using a data set with more than two components, $(x_{1i}, x_{2i}, \dots, y_i)$, and fitting a model equation with multiple dimensions $\hat{y}_i = \beta_0 + \beta_1x_{1i} + \beta_2x_{2i} + \dots$. The definition of $R^2 = 1 - \text{SSE}/\text{SST}$ remains unchanged. However, as more parameters are added, R^2 increases, so to select an optimal model a statistic is required that corrects for the effect of the additional parameters in the model. This statistic is R^2_{adj} that corrects R^2 for the degrees of freedom lost due to the addition of more model parameters, $R^2_{\text{adj}} = 1 - [\text{SSE}/(N - p)]/[\text{SST}/(N - 1)]$, where p is the number of parameters in the model [31].

APPENDIX B: SURVEYS

The questions used in the two surveys asked the students if some study action was performed, how much time was spent in the performance of the action, and the relation between when two actions were performed. An example set of survey questions follows.

Questions asking if an action was performed and how many times an action was performed.

Survey 1—Question 1.—By the time you took Exam 1, how many times had you read through the assigned chapters in the course guide?

Select One of the Following:

- (a) 0
- (b) 1
- (c) 2
- (d) 3
- (e) more. How many? _____

Survey 2—Question 1. This question asks about lecture notes:

- (a) Do you take lecture notes? (Circle One) (a) Always, (b) Sometimes, (c) Rarely, (d) Never.
- (b) Do you use the notes while you work the homework? (Circle One) Yes or No.
- (c) Did you review your lecture notes for the exam? (Circle One) Yes or No.

Survey 2—Question 2 This question asks you about your use of the online homework solutions.

- (a) Did you review any of the online homework solutions for the last test? (Circle One) Yes or No.
- (b) If yes, how many of the six homework solutions covering material on test 3 did you review online as you studied for the test? (Circle One) 6 5 4 3 2 1 0

Survey 2—Question 11. There were three practice tests for Exam 3.

- (a) How many practice tests did you read? (Circle One) 3 2 1 0
- (b) How many practice tests did you work? (Circle One) 3 2 1 0
- (c) How many practice tests solutions did you read? (Circle One) 3 2 1 0

Questions asking how much time was spent performing an action.

Survey 1—Question 3 How much time (in hours) do you spend on the average homework set in UPII?

Survey 2—Question 12 This question asks you about how time was spent studying for Exam 3 in hours. The sum of the time spent on each action should add to the total time spent on the exam.

- (a) time spent reading course guide chapters _____
 - (b) time spent working examples in the course guide _____
 - (c) time spent reading other textbooks or study guides _____
 - (d) time spent reading or working practice tests _____
 - (e) time spent reviewing homework solutions _____
 - (f) time spent reviewing lecture notes _____
 - (g) time spent reworking homework problems _____
 - (h) time spent asking questions or getting help _____
 - (i) time spent on other methods of studying _____
 - (j) total time spent studying for Exam 3 _____
- Questions asking the relative timing of two actions.*

Survey 1—Question 2 When do you first do the assigned reading?

Select One of the Following:

- (a) before lecture
- (b) while doing the self-test part of the homework
- (c) while working the homework
- (d) before the exam
- (e) never
- (f) other—Explain

- [1] A. W. Chickering and Z. F. Gamson, Seven principles for good practice in undergraduate education, *AAHE Bulletin* **39**, 3 (1987).
- [2] George D. Kuh, Assessing what really matters to student learning, *Change* **33**, 10 (2001).
- [3] National Survey of Student Engagement, *Major differences: Examining student engagement by field of study annual*, Indiana University Center for Postsecondary Research, Bloomington, IN (2010).
- [4] Rosanne Di Stefano, Physics teaching and time management, *Phys. Teach.* **36**, 350 (1998).
- [5] Lawrence A. Coleman, Donald F. Holcomb, and John S. Rigden, The introductory university physics project 1987-1995: What has it accomplished? *Am. J. Phys.* **66**, 124 (1998).
- [6] Wayne W. Welch and Robert G. Bridgham, Physics achievement gains as a function of teaching duration, *School Sci. Math.* **68**, 449 (1968).
- [7] Leonard Springer, Mary Elizabeth Stanne, and Samuel S. Donovan, Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis, *Rev. Educ. Res.* **69**, 21 (1999).
- [8] Wayne C. Fredrick and Herbert J. Walberg, Learning as a function of time, *J. Educ. Res.* **73**, 183 (1980).
- [9] Deanna M. Spanjers, Matthew K. Burns, and Angela R. Wagner, Systematic direct observation of time on task as a measure of student engagement, *Assessments for Effective Intervention* **33**, 120 (2008).
- [10] James A. Kulik and Chen-Lin C. Kulik, Review of recent research literature on computer-based instruction, *Contemp. Educ. Psychol.* **12**, 222 (1987).
- [11] Yiping Lou, Philip C. Abrami, and Sylvia d'Apollonia, Small group and individual learning with technology: A meta-analysis, *Rev. Educ. Res.* **71**, 449 (2001).
- [12] Pu-Shih Daniel Chen, Amber D. Lambert, and Kevin R. Guidry, Engaging online learners: The impact of web-based learning technology on college student engagement, *Comput. Educ.* **54**, 1222 (2010).
- [13] Reid Bates and Samer Khasawneh, Self-efficacy and college students' perceptions and use of online learning systems, *Comput. Hum. Behav.* **23**, 175 (2007).
- [14] Avram Goldstein, Does homework help? A review of research, *The Elementary School Journal* **60**, 212 (1960).
- [15] Rosanne A. Paschal, Thomas Weinstein, and Herbert J. Walberg, The effects of homework on learning: A quantitative synthesis, *J. Educ. Res.* **78**, 97 (1984).
- [16] Alexander Renkl, Learning from worked-out examples: A study on individual differences, *Cogn. Sci.* **21**, 1 (1997).
- [17] Barbara Means, Yukie Toyama, Robert Murphy, Marianne Bakia, and Karla 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, DC, 2010).
- [18] Dov Liberman and H. T. Hudson, Correlation between logical abilities and success in physics, *Am. J. Phys.* **47**, 784 (1979).
- [19] H. T. Hudson and Dov 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).
- [20] Warren Wollman and Frances Lawrenz, Identifying potential "dropouts" from college physics classes, *J. Res. Sci. Teach.* **21**, 385 (1984).
- [21] Ibrahim Abou Halloun and David Hestenes, The initial knowledge state of college physics students, *Am. J. Phys.* **53**, 1043 (1985).
- [22] David 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).
- [23] Audrey B. Champagne, E. Klopfer, and John H. Anderson, Factors affecting the learning of classical mechanics, *Am. J. Phys.* **48**, 1074 (1980).
- [24] Philip M. Sadler and Robert H. Tai, Success in introductory college physics: The role of high school preparation, *Sci. Educ.* **85**, 111 (2001).
- [25] David P. Maloney, Thomas L. O'Kuma, Curtis Hieggelke, and Alan Van Huevelen, Surveying students' conceptual knowledge of electricity and magnetism, *Am. J. Phys.* **69**, S12 (2001).
- [26] Joan Laing, Richard Sawyer, and Julie Nobel, Accuracy of self-reported activities and accomplishments of college-bound students, *J. Coll. Student Dev.* **29**, 362 (1988).
- [27] George D. Kuh, *The National Survey of Student Engagement: Conceptual Framework and Overview of Psychometric Properties* (Indiana University, Center for Postsecondary Research, Bloomington, IN, 2011), http://nsse.iub.edu/pdf/psychometric_framework_2002.pdf, accessed May 11, 2011.
- [28] David J. Palazzo, Young-Jin Lee, Rasil Warnakulasooriya, and David E. Pritchard, Patterns, correlates, and reduction of homework copying, *Phys. Rev. ST Phys. Educ. Res.* **6**, 010104 (2010).
- [29] Richard R. Hake, Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses, *Am. J. Phys.* **66**, 64 (1998).
- [30] Jay L. Devore, *Probability and Statistics* (Thomson Brookes-Cole, Toronto, 2004), 6th ed.
- [31] Douglas B. Montgomery, Elizabeth A. Peck, and G. Geoffrey Vining, *Introduction to Linear Regression Analysis* (John Wiley and Sons, New York, 2001) 3rd ed.