Teasing Out the Effect of Tutorials via Multiple Regression

Stephanie V. Chasteen

Science Education Initiative and Department of Physics, University of Colorado, Boulder, CO 80309, USA

Abstract. We transformed an upper-division physics course using a variety of elements, including homework help sessions, tutorials, clicker questions with peer instruction, and explicit learning goals. Overall, the course transformations improved student learning, as measured by our conceptual assessment. Since these transformations were multi-faceted, we would like to understand the impact of individual course elements. Attendance at tutorials and homework help sessions was optional, and occurred outside the class environment. In order to identify the impact of these optional out-of-class sessions, given self-selection effects in student attendance, we performed a multiple regression analysis. Even when background variables are taken into account, tutorial attendance is positively correlated with student conceptual understanding of the material – though not with performance on course exams. Lecture attendance, which includes exposure to clicker questions and peer instruction, did not achieve the same impacts.

Keywords: physics education research, course reform, electricity and magnetism, assessment
PACS: 01.30.Ib, 01.40.Di, 01.40.Fk, 01.40.G-, 01.40.gb

INTRODUCTION

An increasing number of physics education research groups are turning their attention to the upper-division courses such as thermodynamics, quantum mechanics, and electricity and magnetism[1]. Among the variety of instructional approaches used in these course transformations are conceptually-focused small-group activities, such as tutorials. In a complete course overhaul, however, many things are changed at once. It is therefore difficult to discern the effect of individual elements; yet this information is important in order to identify future directions for fruitful research and development. The current paper focuses on the impact of the tutorials developed as part of our course transformations in junior level E&M.

E&M COURSE TRANSFORMATIONS

Over 4 years, we have modified the first semester of a two-semester junior-level sequence in electro- and magneto-statics (E&M1), typically taken in the fall of the junior year. Around 25-50 students enroll in a given semester of E&M1. The transformed course includes explicit consensus learning goals, modified homework, traditional lectures with interactive elements such as clickers and peer instruction, homework help sessions, and optional tutorials[2]. The transformed course elements have been used in 5 courses at the University of Colorado (CU) [3] as well as by outside institutions. We followed the course transformation model[4] developed by the Science Education Initiative (SEI)[5].

To assess the relative success of these transformations, and to document student difficulties, we developed a post-test. The Colorado Upper-Division Electrostatics (CUE)[6] is an assessment consisting of 17 open-ended questions, showing high inter-rater reliability and validity. The CUE tests a variety of skills, including students' ability to choose a problem-solving method, sketch electric fields, graph electric fields and potentials, and explain the physics and mathematics in common problems. A pre-test was developed from a subset of the questions on the CUE.

We have previously shown that students in courses using the transformed materials score higher on the CUE post-test, on average, than those using traditional lecture-based instruction [3,4], and that these results hold regardless of student background. Students in 16 courses received the CUE as an in-class post-test during the last week of class: 8 transformed courses (5 at CU) and 8 traditionally-taught courses (2 at CU). Taking each student as a data point (N=488), the average CUE score is higher in the transformed courses (58.2 ± 1.4%, 8 courses, 189 students) than in standard courses (44.6 ± 1.6%, 8 courses, 299 students, $p<0.001$). However, this improvement is undoubtedly due to multiple factors. In the rest of the paper we examine the effect of the optional tutorial sessions. The study includes 5 CU courses (N=205)[7] using the tutorials and other transformed materials.

THE TUTORIALS

A series of optional weekly tutorials was developed and refined over two years, with the later addition of tutorial pre-tests (“preflights"). Tutorials were designed to reinforce topics presented in lecture, expand on these topics, and prepare students for the upcoming homework. Student attendance was optional but acceptably high (30-44%; average 38%), and students worked in groups of 3-5 to complete a conceptually-focused worksheet on the material. Optional homework help sessions were also offered, in which students worked on homework in groups with the assistance of the instructor.

Both homework help sessions and tutorials were geared to help students develop metacognitive strategies and communication skills, as well as to allow the instructor to model effective problem-solving strategies. Both types of sessions also offered a valuable chance for students and instructors to interact, providing instructors with insights into student thinking. Instructors were positive about their experience with the tutorials[3].

When asked to rate course elements as useful for their learning, students indicated that lecture, clicker questions, and tutorials were most helpful. Tutorials were rated highly on several measures, and many students commented positively on the tutorials: “I really liked the Friday tutorials. They were (generally) fun, interesting, and a good jump-start to keep me excited over the weekend. Also, I learned a lot.”

In addition to student self-reported learning value and enjoyment, we wanted to determine whether tutorial attendance affected student outcomes. Across all five courses, we find that, the more tutorials attended, the higher a student’s exam scores (Pearson’s r=0.27), course grades (r=0.36), and CUE scores (r=0.26) (all $p<0.01$, N=158-198). Does this mean that tutorials have a positive impact on student performance?

We must be careful in drawing such conclusions, since the optional nature of tutorials results in self-selection effects. Students who go to more than 3 tutorials (out of ~12) tend to be the better students, with higher course grades (3.2 versus 2.4, ±0.1, $p<0.001$), exam z-scores (0.23 vs -0.19, ±0.1, $p<0.001$), lecture attendance (87% vs 73%, $p<0.001$), and CUE post-test score (60.7 vs 50.8, ± 2%). $p<0.001$). Thus, tutorial attendance may be a proxy for pre-existing student variables, such as motivation and ability, rather than an impact measure. We note that the tutorial-attending students did *not* score better than low-attenders on pre-course assessments: The BEMA after introductory physics (60 vs 63) or the CUE pre-test before E&M1 (29 vs 30).

About the Multiple Regression Analysis

We wished to determine the effect of tutorials on student performance when background variables are taken into account (to reduce the effect of covariates). Thus, we performed a multiple regression analysis to identify the predictors of two student outcome measures: Exam z-score and CUE score.

We examined a variety of background variables based on correlation strength: Pre-requisite math courses, GPA in all prior math courses, GPA in prior physics courses, cumulative GPA, CUE pre-test, lecture attendance, and scores on the introductory-level Basic Electricity and Magnetism Assessment (BEMA;8), taken at the end of introductory physics[9]. We model these outcome variables as follows:

$$OUTCOME = b_0 + \left(\sum_{k=1}^{N} b_k \times VAR_k\right) + \left(b_{TUT} \times TUTORIAL\right)$$

where *OUTCOME* is either the CUE post-test score or the z-score of the average of the three midterm exams. Exams differ across courses and have a calculational focus, with some conceptual questions. *TUTORIAL* represents the percent of tutorials attended throughout the term, VAR_k are the background variables that are included in the model, b_k are the coefficients for each term. The value of b_{TUT} is the coefficient for the *TUTORIAL* variable, and gives the relative impact of attending the tutorials on *OUTCOME,* all other factors being equal. Variables are entered into the model manually, and background variables, *VAR*, are entered until a model with a high R^2 and the fewest possible background variables is obtained. Then the variable *TUTORIALS* is added.

Because only some students have BEMA scores, the inclusion of this variable reduces N significantly. The sample of students who have taken the BEMA is also a slightly different population – less likely to have tested out of the introductory physics requirement, for example. Thus, we only include the BEMA as a predictor for students who have BEMA scores, and present those models separately.

Regression Results

Results of the regression are shown in Table 1. We find that the most parsimonious model predicting student scores on the CUE (Model 1A and 1B) includes only their GPA in prior physics courses (PHYS GPA). PHYS GPA alone accounts for 23% of the variance in CUE score. For those students for whom we have BEMA scores from introductory physics courses (Model 2A and 2B), the BEMA is a better predictor of CUE scores than is a student’s GPA in prior physics courses. The BEMA on its own

TABLE 1. Multiple regression models to determine impact of tutorials on CUE and exam scores

Model:	**CUE Model 1A**	**CUE Model 1B (w/ tutorials)**	**CUE Model 2A**	**CUE Model 2B (w/ tutorials)**	**Exam Model 1**	**Exam Model 2**
Population	*All students*	*All students*	*Students with BEMA*	*Students with BEMA*	*All students*	*Students with BEMA*
Model statistics						
N	156	156	87	87	192	103
Multiple R^2	0.23	0.26	0.40	0.46	0.46	0.60
F statistic	47.24	$27.08^{\dagger}$	580.8	$36.77^{\dagger\dagger}$	166.93	156.3
Residual std. error	12.26	15.04	13.01	12.41	0.77	0.66
Predictors	b_k	b_k	b_k	b_k	b_k	b_k
PHYS GPA	0.48**	0.45**			0.68**	0.78**
BEMA			0.64**	0.63**		
Tutorials		0.17*		0.24**		

Table 1. Multiple regression statistics: The F-statistic is large if the model's predictive capability is large relative to background variables and error. The residual standard error measures the amount of variance unaccounted for by the model. R^2 is the proportion of variability that is accounted for by the model. All F statistics are significant at $p<0.0001$ value. Coefficients reported are significant at the $p<0.05$ (*) and $p<0.01$ (**) level; if a coefficient is not reported, then it did not enter into the model as a significant predictor. The y-intercept (b_0) is insignificant for all models, and thus is not reported. Significant differences from the previous listed model, as determined by the F-test, is designated by †, $p<0.05$ and †† $p<0.01$.

accounts for 40% of the variance in CUE score. All other background variables were non-significant: That is, their variance was accounted for by the inclusion of PHYS GPA or BEMA.

Regardless of whether a student took the BEMA, the addition of tutorial attendance as a predictor significantly improves the model. This can be seen by the increase of R^2 from Model 1A to 1B, and Model 2A to 2B (see Table 1). The effect of tutorials on CUE scores is roughly one-third that of either PHYS GPA or BEMA. This indicates that tutorial attendance does provide some improvement in performance on conceptual assessments, even when background performance is taken into account. We find the same results regardless of whether tutorial attendance is measured continuously or as a binned variable.

The same is not true for the traditional exams, however: BEMA scores and tutorial attendance did not enter into the model as significant predictors for the exam z-scores (i.e., the difference of a student exam score from the course mean). The PHYS GPA variable alone predicts 46% of the variance in student exam z-scores for the student population as a whole. Thus, it appears that conceptual understanding (as measured by BEMA performance or the experience gained in tutorials) does not strongly affect students' ability on these calculation-focused assessments. While one would hope that the conceptual framework provided by tutorials would enhance student performance on course exams, similar lacks of correspondence between calculational and conceptual performance has been seen elsewhere [8,9]. The conceptual focus afforded by tutorials and the course approach as a whole at least does not *harm* students' calculational skills, as measured by common traditional exam problems[4]. In later semesters we also gave students *conceptually*-focused pre-post quizzes, targeted at the tutorial material. Notably, we also see no learning gains on this post-quiz, given 1-5 weeks after the tutorial.

Lecture and Homework Help Sessions

The other optional, out-of-class activity was the homework help sessions. Attendance at these sessions was only recorded for two out of the five transformed courses at CU. On average, most students (86%) attended at least one help session, but attendance varied widely. On average, a given student attended 40% of the sessions. We consider homework score to be the important outcome variable from these sessions, and performed a multiple regression on the homework scores for the N=61 students for whom we have homework help session attendance data. We find that PHYS GPA is a significant predictor of homework score ($R^2 = 0.41$), and that the percent of help sessions attended improves the model (new model $R^2 = 0.58$).

Does lecture attendance help student learning too? Lecture attendance (as gauged by the presence/absence of a student response to clicker questions in a particular lecture) is moderately correlated with post-test score on the CUE ($r=0.199$, $p<0.05$, N=161), and more strongly correlated with traditional measures such as course grade ($r=0.35$, $p<0.001$, N=201) and average course exam score ($r=0.29$, $p<0.001$, N=200). However, this appears to be mostly due to a self-selection effect: In the linear regression models, we found that lecture attendance was *not* a significant predictor of student scores on the CUE or course exams when grades in prior physics courses were taken into account, but a low spread on this variable makes it difficult to discern effects of lecture attendance in transformed courses (i.e., most students attend most lectures). Additionally, attendance is higher in transformed versus traditional courses, but low N prevents us from performing multiple regression in the traditional courses.

DISCUSSION AND CONCLUSIONS

We find that, when controlling for background variables such as grades in prior courses, the best predictor of student success on our conceptual exam (CUE) and traditional course exams is the student's GPA in previous physics courses. For those students with BEMA scores, success on the BEMA is a very strong predictor of success on the CUE. This may be related to student ability, and/or to the fact that the BEMA and CUE both provide measures of student motivation (to work hard on an ungraded exam).

Regardless of whether a student had a BEMA score or not, tutorial attendance was a significant predictor of success on the CUE. This correlation suggests that the tutorial experience provided students with conceptual understanding and reasoning skills measured on the CUE (assuming that covariates such as motivation are removed by the background variables). The tutorials achieved these positive results despite the fact that their development was guided by only very preliminary research into student thinking at this level. Thus, while the aim is to target tutorials towards researched student difficulties, simply providing an opportunity for students to engage in such activities – making sense of course material, working with their peers, and interacting with the instructor – may have intrinsic benefits. These positive findings provide support for the use of such focused activities – and perhaps for providing such activities in-class, for the benefit of all students. We are interested in comparing such hour-long worksheets to shorter (e.g., 15 minute) in-class activities that are easier to integrate and target to single concepts.

However, tutorial attendance did not predict student success on traditional course exams (and thus, by proxy, student success in the course), and student performance does not improve after the tutorial on conceptual pre/post-quizzes on the material. One interpretation of these results is that additional research and development is needed to align the tutorial outcomes with factors influencing a student's course grade, and we ought to focus the tutorials on specific challenging concepts. Or, the value of tutorials may lie elsewhere. There are many positive outcomes of tutorials for students, such as conceptual understanding, positive attitudes, and communication skills practice. CUE scores may be affected by tutorials (whereas pre/post quizzes are not) because students are not learning specific *concepts* but rather the *habits of mind* necessary to figure out new problems. During tutorials, students have a chance to act like physicists (debating and reasoning) in a way not assessed or supported elsewhere, potentially affecting the class culture. Students are not the only ones affected by the tutorials; faculty co-teach the sessions, and gain a valuable window into student thinking. According to one seasoned PER instructor: "The asset that came as a surprise to me, because I thought I knew it all, was how valuable the feedback is that I'm getting in the tutorials… The tutorials seem to reveal to me the students' thinking, or lack of it, in a way that watching them struggle with the homework doesn't…. It's just eye-opening."

Thus, tutorials provide multiple advantages (both quantitative and qualitative) be to be considered when weighing the time and cost of development and facilitation.

ACKNOWLEDGMENTS

This work is funded by the SEI, CU-Boulder and NSF-CCLI grant #0737118. I'd like to thank Lauren Kost for many helpful comments on the multiple regression analysis, and Rachel Pepper, Steven Pollock, Katherine Perkins, and anonymous reviewers for valuable comments on the analysis and the paper.

REFERENCES

1. See, for example, S.V. Chasteen and S. J. Pollock, *PERC Proceedings 2009*, **1179**, 109-112 (2009); and B. Ambrose, Am. J. Phys., **72**(4), 453-459 (2004); and references within.
2. The full set of course materials are available at www.colorado.edu/sei/departments/physics_3310.htm
3. S.V. Chasteen, R.E. Pepper, S.J. Pollock and K. K. Perkins, *PERC Proceedings 2011,* (submitted).
4. S. V. Chasteen, K. K. Perkins, P. Beale, S.J. Pollock and C. E. Wieman, *J. College Sci. Teach,* **March/April,** 70-76 (2011).
5. See http://colorado.edu/sei
6. S. V. Chasteen and S.J. Pollock, *PERC Proceedings 2009,* AIP Conference Proceedings, **1179,** 109-112 (2009).
7. Described in more detail in S. Chasteen, R. Pepper, S. Pollock and K. Perkins, *PERC Proceedings 2011* (in press).
8. L. Ding, R. Chabay, B. Sherwood, *Phys. Rev. ST Physics Ed. Research*, **2**(1), 1-7 (2006).
9. S.J. Pollock, *Phys. Rev. ST Physics Ed. Research,* **5(**2), 020110-020117 (2009).
10. C. Hoellwarth, M. J. Moelter and R. D. Knight, " *Am. J. Phys.,* **73** (5), 459-462 (2005)