The effects of a redesign on students' attitudes and evaluations
for an introductory calculus-based physics lab

Saif M. Ali and Brian D. Thoms

Department of Physics & Astronomy, Georgia State University, 25 Park Place, Atlanta, Georgia, 30303

This study assesses student reactions to a redesign of an introductory calculus-based physics lab. A three-hour verification-style lab was replaced with a one-hour tutorial and two-hour inquiry-based lab. After confirming that the redesign improved students' performance in a separate study, this work evaluates students' opinions to determine if they were "buying in" to the redesigned process. When the lab concluded, the students from both labs were asked a series of Likert-scale and free-response questions. The responses were analyzed using a coding rubric to determine students' thoughts on the tutorial and lab. This revealed what aspects of the lab the students believed most positively and negatively affected their learning. It was observed that students had an overwhelmingly negative take on the new lab's tutorial section. Interestingly, tutorials were also among the most frequent positive comments. This finding, along with others, was scrutinized to check for any trends that allow us to make hypotheses and significant observations.

I. INTRODUCTION

In the Fall of 2014, Georgia State University redesigned the laboratory portion of the two-semester, introductory, calculus-based physics course. The previous, traditional lab was three hours in length and revolved around experiments led by a single graduate Teaching Assistant (TA) where verification of theory was emphasized. In the redesign, the lab was split into two back-to-back sections. The first section was an hour-long tutorial led by a trained, undergraduate Learning Assistant (LA) followed immediately by a two-hour experiment led by a TA. The tutorial section used published materials from the University of Washington [1] which emphasize conceptual learning. The two-hour section was an inquiry-based experiment where the students were presented with a question on the topic at hand, asked to make predictions based on previous conceptions and coverage of material in lecture and textbook, run an experiment, and then draw conclusions based on the data gathered. The labs were designed so that the results could be compared with predictions made, and the students would lastly be asked to communicate their findings in short-answer form.

The main goals of the redesign were to make the lab more interesting, have it correspond well with the lecture, improve the writing of the lab manuals, increase conceptual understanding via the tutorials, and to improve student attitude towards the lab. While it would be easier to interpret results from a single change such as introducing tutorials and learning assistants or introducing inquiry-based experiments, the constraint of fitting into the existing three-hour time block required experiments to be modified significantly in order to accommodate time for tutorials. This study attempts to understand the effects on students attitudes and their causes as far as possible given this constraint.

In a previous study done at Georgia State University [2], the author's group found that for the first semester course the Force Concept Inventory (FCI) scores increased after the redesign, as shown in Table I, indicating increased conceptual understanding regarding forces. Using the Hake gain [3], $g = \frac{Post - Pre}{100\% - Pre}$, average normalized student gain, <g>, increased from 0.30 in the last year before the redesign to 0.40 in the first year of the redesign, a difference that was found to be statistically significant at the p<0.01 level.

Grades also increased with the introduction of the new lab design with pass (ABC) rates increasing from 86% to 91% (significant at p<0.0001) as seen in Table II. While the increases in conceptual understanding and student success meet many of the goals of the instructors, this study investigates the success of the redesign from the student perspective by analyzing end-of-semester comments before and after the redesign.

TABLE I. Comparison of FCI Data

	N	FCI Pre (%)	FCI Post (%)	<g>
Conceptual Understanding				
Traditional	218	36.3	54.8	0.30
Redesigned	227	36.6	61.0	0.40

TABLE II. Comparison of Course Grade Point Average and Success and Withdrawal Rates

	N	GPA	ABC	DWF	W
Grades and Success Rates					
Traditional	421	3.12	86.0%	14.0%	8.8%
Redesigned	521	3.29	91.2%	8.8%	4.6%

II. SURVEY

To ascertain students' opinions on various aspects of the lab, a survey consisting of seven statements with Likert scale responses ranging from Strongly Agree to Strongly Disagree and three free-response questions was given in lab at the end of the semester. For the Likert scale questions, students were asked to rate their level of agreement to statements regarding the lab's helpfulness in understanding physics concepts, how interesting it was, the correspondence of the lab with the course, writing quality of the manuals, lab difficulty, and instructor's preparedness and quality of teaching. Although these survey questions were not validated, they were drawn from the most common student comments prior to the redesign. The number of occurrences for each possible response in a Likert scale was summed for each statement. The open-ended questions on the survey asked students to identify the most useful part of the lab portion of the course, the part that needs the most improvement, and any additional comments they had.

Data used in this study were collected in the first course of the two-semester sequence. Opinions on the traditional lab were collected in the Spring 2014 and Summer 2014 semesters while the data for the redesigned lab were collected in the Fall 2014, Spring 2015, Summer 2015, and Fall 2015 semesters. In total, there were 699 students who responded to the Likert scale statements and 1244 answers to the free-response questions.

III. LIKERT SCALE QUESTIONS

A chi-square test was applied to determine significant differences before and after the redesign in student response patterns of the seven Likert scale questions. Only the question regarding the correspondence of the lecture and lab portions of the course showed a significant difference. Before the redesign students 25% of students agreed and 16% strongly agreed that the lab and lecture corresponded well while 24% and 12% disagreed and strongly disagreed, respectively. After the redesign, strongly agree and agree increased to 25% and 47%, respectively, while strongly disagree and disagree decreased to 2% and 9%, respectively. The answer distributions were found to be statistically different at the $p < 0.00001$ level using a chi-square test. One of the goals of the redesign was to create a lab schedule that more realistically matched the pace of instructors, so in this case student opinions indicate some level of success in the redesign in this area.

The largest change (other than correspondence with lecture) was to the overall statement regarding the benefit of the lab in learning concepts. Strongly agree responses increased for that question from 26% to 33%. While this difference was not statistically significant, it is interesting given the fact that FCI results show significantly improved student conceptual understanding.

IV. FREE RESPONSES AND CODING SCHEME

Three free-response questions were given in the survey asking students the most useful part of the lab portion of the course, the part that needs to be improved, and any other comments they would like to give. The free responses were analyzed by developing a coding system. The first step was to identify the recurring topics and create a one- or two-letter code for each. For example, responses discussing the tutorials were coded with T. More than one code could be applied to a response. Also, since students were prompted specifically for positive and negative comments, a plus or minus sign was to be added to every code to designate the tone of the response. For example, if a student said, "The lab was great, but the equipment made it difficult sometimes", then it would be coded with "LE+" which are for general comments about the labs and experiments and "E-" which represents lab equipment.

The first version of the codes was based on a sample set of responses taken from a summer semester. The coding scheme was very specific at first, in an attempt to have all the responses covered. The codes were tightened by grouping together closely related codes. Next, two researchers applied the codes to a sample of responses and compared results. Differences in coding were discussed to clarify descriptions or combine codes. After three rounds an inter-rater reliability of 85% was achieved on a sample of more than 10% of the responses. The final code list is shown in Table III.

TABLE III. Coding Scheme

Code	Association	Code	Association
HO	Hands-On	NB	Notebook
CU	Conceptual Understanding	QD	Question/Direction Difficulty
GW	Group Work	Q	Quizzes
TA	Teaching Assistant	G	Grading
T	Tutorials	PL	Pre-Lab (Old Lab)
LM	Lab Manuals	M	Mathematics
ON	On-Pace with Lecture	V	Visualization
E	Equipment	RW	Real-World Environment
LE	Labs/Experiments		
HW	Homework	TI	Time Allocation
LA	Learning Assistant	O	Other

Every free response was coded with this final coding scheme. The occurrences of the codes were counted in order to compare student opinions on the traditional and redesigned lab. Since the tone of the codes is important in interpreting the results, positive and negative codes were counted separately. The occurrences of the codes were then summed for the traditional and redesigned labs and the percentage of the total number of positive or negative codes on each topic was calculated to allow a direct comparison.

V. COMMENT TRENDS

To identify significant trends in free-response answers the change in code percentages between traditional and redesigned labs were calculated as shown in Table IV. The largest change is for the tutorial code (T), which did not occur for the traditional labs since there were no tutorials, but comprised about 6% of the positive codes and 20% of the negative codes for the redesigned labs. Since this is by far the dominant change, we will discuss this in detail in the next section. Increases of 4% and 5%, respectively, are observed for codes related to lab homework and allocation of time within the lab. Both of these appear related to tutorials and will also be discussed in the next section.

Other than the tutorial comments, changes in code occurrences ranged from about -10% to +5%. One significant change was a 10% drop in the occurrence of negative comments for the lab being on time with the lecture. This agrees with the change observed on that topic in the Likert scale questions, but in addition indicates that this is important to students. A decrease of 6.2% was observed in the occurrence of the code counting negative comments regarding lab manuals which were a focus of the redesign. A 5.2% increase was found in the occurrence of positive lab and experiment (LE) comments, but occurrences of the conceptual understanding (CU) code showed only small changes. Although, in Likert scale questions, student agreement that the lab helped them learn physics concepts increased modestly. The significant change in the number of positive comments of the lab despite a relatively small change in overall satisfaction with the experiments may indicate that this aspect is important to students and that the redesign removed elements that were particularly dissatisfying for some students. It is not clear whether the students' reactions are due to the type of equipment, learning materials, or experiment format. All were changed in the redesign, since changing only one variable in isolation was not possible.

TABLE IV. Percent Changes in Comment Occurrences from Traditional to Redesigned Lab

	Comments (%)			Comments (%)	
Code	Negative	Positive	Code	Negative	Positive
HO	0.1	-1.3	NB	-1.7	-0.3
CU	-0.3	0.2	QD	-1.2	-1.3
GW	0.1	1.5	Q	-5.1	-0.3
TA	-2.5	0.4	G	-2.6	-0.3
T	20.3	5.7	PL	-0.9	-2.4
LM	-6.2	-0.3	M	-0.5	-3.1
ON	-10.4	1.1	V	0.2	-1.0
E	0.5	0.3	RW	0.0	0.9
LE	0.3	5.2	TI	4.6	0.0
HW	4.2	0.2	O	-0.7	-1.3
LA	0.7	2.1			

VI. TUTORIALS

The most significant change observed was for the occurrence of the tutorial (T) code. Twenty percent of the negative comments involved the tutorials. To better determine the reasons for this student discontent, we counted the co-occurrence of other codes in the same responses as negative tutorial comments, which we refer to as sub-codes. For example, one student stated the "Tutorials are not written very clearly", so this would be coded with "T" and "QD" which denotes when a student mentions the quality or difficulty of the directions of the lab or tutorial.
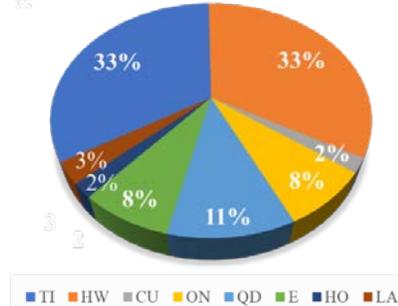

FIG 1. Percentage breakdown of sub-codes (codes co-occurring with T).

As shown in Fig. 1, there were eight sub-codes that were attached to the tutorial comments. These included HW, CU, QD, E, HO, LA, and TI. HW and TI each accounted for one third of the sub-codes for tutorial comments. At a total of 34 comments for each topic, it is obvious that homework and time allocation were among the main issues the students had with the tutorials.

For the traditional labs, homework involved writing lab reports submitted before the next lab. For redesigned labs, students recorded data and answered questions while completing the experiment and handed in their work at the end of the lab period so they did not have to write a formal lab report, but they did have tutorial homework using the homework manual associated with the UW tutorials [1]. Comments indicate that some students felt that the tutorial homework was too difficult or they didn't see any benefit from doing it. One student who talked about homework said, "The homework portion needs to be removed. We already have homework for this class. It needs to be shorter." Along with the homework for the lecture, having a homework assignment appears to have frustrated some students. Although the time needed to complete the tutorial homework is probably less than the time needed for lab reports in the traditional labs, students appear to resent more having to do the tutorial homework. It may be that tutorials and tutorial homework are not part of what they expect to be doing in the lab portion of the course and seem too close to work done in the lecture portion.

The issue with time allocation the students had mostly had to do with the length of the tutorials. Students were graded in the tutorial portion for being engaged in the tutorial and not for correctness or completeness of their

work. LAs were instructed to end the tutorial after one hour even if students are not finished with the tutorial to leave enough time for the experiment. Many comments said the tutorials ran too long, and a few said that they observed their classmates losing focus before the end of the tutorial session. One student went as far as to state that they wish they could "just spend all the time on the lab itself." While most students did not say to get rid of the tutorials all together, there was many who stated they wanted to move on from the tutorial and start the lab instead of waiting for the hour-long tutorial to be done. Student issues with time allocation may indicate that they see the tutorials as providing less benefit than the lab.

VII. RELATION TO OTHER RESEARCH

In previous work on reactions to tutorials Cruz *et al.* [4] reported that students "complained, and some state that all the tutorials would do is lower their GPA" and that they also "claim that, as they got things wrong during the in-class sessions, their mistakes propagated into their homework." The authors analyzed the test scores of the students with and without tutorials and they found that the students who were doing the tutorials had a higher average grade. The current study also shows that a redesign including tutorials results in higher student grades. In addition, the FCI results reported here show that the improved grades are accompanied by increased conceptual understanding. The students at Georgia State also had a negative perception of the tutorials, but their grades increased despite that. A study by Riegler *et al.* [5] observed the experiences of German students with these same tutorials and found that the students "frequently characterized *Tutorials* as 'annoying', often times because they are 'moving too slowly'. This opinion was echoed by the students at Georgia State, one of which said "shorten tutorials" and another who said the tutorial "drags on at times". Finkelstein and Pollock [6] also report that a sizeable fraction of students do not enjoy the tutorial experience even though they enjoy the group work and show improved conceptual understanding. This tells us that across different classes in different contexts, there are fundamental issues with the tutorials that lead to some student dissatisfaction. The current analysis uses a coding system to quantify student comments and opinions.

VIII. CONCLUSIONS

As shown by the increase in the FCI gains and grades of the students, the redesign resulted in improved student performance. Yet, there were still a significant number of students who did not buy into a major part of the redesign – the tutorials. It is natural to surmise that a lack of acceptance and increased dissatisfaction will limit the benefit students would get from the redesigned lab. Therefore anything that can be done to improve student buy-in or increase the value that students place on the tutorials will lead to increased effectiveness.

Clearly, a number of students don't see tutorials as the most valuable or effective use of their time and energy. One possible reason that students were not on-board could be that the goals of the lab were not communicated to them well enough. The separation of teaching responsibilities between the LA (tutorials) and TA (experiment) may contribute to a less than coherent message about the relationship between the two parts of the lab. Lecture instructors may be able to do more to reinforce the value of conceptual understanding and of the tutorials as a way to achieve it.

It is also clear that tutorial homework is not valued highly by the students. More work is needed to determine the effectiveness of the tutorial homework in this lab design. It is possible that less homework would be valued more and therefore be more effective. It's also possible that with less tutorial homework the students would have less incentive to engage deeply in the tutorials themselves.

Overall, this study suggests that an integration of a tutorial and experiment in a single lab experience is effective for student learning but creates challenges for getting student buy-in for the tutorial portion of the lab.

[1] L. C. McDermott and P. S. Schaffer, Tutorials in Introductory Physics (Prentice-Hall, Upper Saddle River, NJ, 2002).

[2] Thoms, B., Topdemir, Z., Trusty, D., Oncul, E., Doluweera, S., & Von Korff, J. Bulletin of the American Physical Society, Volume 61, Number 1 (2016).

[3] Hake, R. R. American journal of Physics 66 (1), 64-74 (1998).

[4] Cruz, É., O'Shea, B., Schaffenberger, W., Wolf, S., & Kortemeyer, G. The Physics Teacher 48(7), 453-457 (2010).

[5] Riegler, P., Simon, A., Prochaska, M., Kautz, C., Bierwirth, R., Hagendorf, S., & Kortemeyer, G. Physics Education, 51(6), 065014 (2016).

[6] Finkelstein, Noah D., and Steven J. Pollock. Physical Review Special Topics-Physics Education Research 1.1 (2005): 010101.