

Impact on students' views of experimental physics from a large introductory physics lab course

Nidhal Sulaiman^{1,2}, Benjamin Pollard¹, and H. J. Lewandowski¹

¹*JILA, National Institute of Standards and Technology and the University of Colorado, Boulder, CO 80309, USA
Department of Physics, University of Colorado Boulder, Boulder, CO 80309, USA and*

²*Department of physics, Sultan Qaboos University, P.O. Box 36, AlKhod, Oman*

Introductory physics lab courses aim to have students gain a wide variety of skills and knowledge, including developing views of the nature of experimental physics that are aligned with common expert views. The large introductory lab course (~ 700 students) at the University of Colorado Boulder has been recently transformed to explicitly address this goal among others. To measure the level of success in reaching this goal, we used an established assessment instrument, the Colorado Learning Attitudes about Science Survey for Experimental Physics (E-CLASS), which probes students' views and expectations of experimental physics. We collected students' responses to E-CLASS during three semesters before, and four semesters after, the course transformation. We observe statistically significant differences between the before and after transformation post-test scores of the (i) overall E-CLASS survey and (ii) some individual E-CLASS items, especially those closely related to specific course learning outcomes.

I. INTRODUCTION

Introductory Physics labs can have a variety of desired learning outcomes from enhancing understanding of physics concepts, to increasing hands-on technical skills, to developing professionally aligned views of experimental physics. The American Association of Physics Teachers (AAPT) has produced an extensive guide for possible goals for both introductory labs and labs beyond the first year of college titled “Recommendations for the Undergraduate Physics Laboratory Curriculum” [1]. These recommendations focus on six major themes, namely, constructing knowledge, modeling, designing experiments, developing technical and practical laboratory skills, analyzing and visualizing data, and communicating physics. Since the publication of these recommendations, many institutions have worked to modify their lab classes to align with these goals [2–7].

At the University of Colorado Boulder (CU), the introductory lab course for physical science and engineering majors has been completely transformed to better align with relevant learning goals for students in these departments and the AAPT guidelines [1]. This transformation process began with establishing consensus learning goals, which focused on aspects of affect, measurement uncertainty, and views of the nature of experimental physics, among others and is described in more detail in Ref. [8]

A critical component of lab course transformations, including the one at CU, is careful and repeated measurements of the impact of the lab course on student learning. Some aspects of one of these goals, developing expert views on the nature of experimental physics, can be measured using a research-based assessment tool called the Colorado Learning Attitudes about Science Survey for Experimental Physics (E-CLASS) [9]. This tool probes students’ epistemology and expectations of experimental physics in the context of a lab class by asking about students’ agreement with 30 statements in a Likert-style format.

To assess the effectiveness of the CU lab transformation, we administered the E-CLASS to students both before and after the transformation. We use these data to measure the impact of the course transformation along one particular dimension. Here, we present results of analysis of these data to measure impact on both the overall E-CLASS score and on scores item-by-item. By looking at the change in student responses to each item, in combination with the established learning goals and curriculum, we can better understand how the lab activities and structure are impacting student learning.

Compared to previous work examining the impact on E-CLASS scores in this class [10], this study involves a significantly larger data set, where E-CLASS data were collected over seven semesters, three of which were before transforming the lab course (BT) and four were after transformation (AT), and include data from different instructors. In total, the data set represents 3247 students. Additionally, here, we use ANCOVA to examine student responses to all 30 post-test E-CLASS items, while controlling for their pre-test scores.

II. BACKGROUND

Course Context. The introductory lab course at CU is a one-credit course that covers mechanics and electricity and magnetism content. Students attend six lectures throughout the semester and a two-hour weekly lab section. Typically, enrollments are between 600 and 700 students who are mostly first- and second-year undergraduates and are concurrently taking the second semester physics lecture course.

Before the transformation, the learning goals of the course were focused on exposing students to a variety of simple lab equipment, teaching students error propagation, writing a properly formatted lab report, and training students to use the Mathematica software package. Students performed six experiments in total throughout the semester, with each experiment spanning two weeks. Students would take data in the first week and perform data analysis and write a report in the second. Additionally, students completed six homework assignments on error analysis and propagation. Most of the points earned for the final grade in the class came from lab reports.

For the transformation, lab activities and lectures were created to address new learning goals [11]. These goals included aligning student beliefs with expert views around the nature of experimental physics, having a positive attitude about the course and towards experimental physics, making presentation quality graphs of models and corresponding data, and developing set-like reasoning around measurement uncertainty [12].

The transformed course includes pre-lab activities and 12 two-hour lab activities which are grouped into 4 modules (skill building, mechanics, electronics, and optics) [13]. As in the previous course, students work in pairs. However, each student keeps their own lab notebook on a tablet PC using OneNote and uploads a PDF version to the learning management system at the end of the lab session. Most of the points earned for a grade in the class came from lab notebooks. Lab reports are no longer part of the course. Activities are designed so that students are prompted to compare their predictions and findings with other groups. To further increase students’ understanding of measurement uncertainty and related statistics, students are sometimes asked to combine their data with those collected by all other groups and reflect on the data set as a whole.

E-CLASS. E-CLASS is attitudinal and beliefs survey that is focused on experimental physics and lab courses. The survey is administered at the beginning and end of a semester using an online automated system [9, 14]. For both pre- and post-test surveys, students are asked to rate their level of agreement to 30 statements such as, “Working in a group is an important part of doing physics experiments.” Evidence of validation of E-CLASS has been presented by others [15] and many studies have been done using data from individual institutions [9] and data collected across the US [16, 17]

III. METHODS

Data Collection. We collected responses to the E-CLASS from students in the class both before the course transformation (BT), from Fall 2016 through Fall 2017 (3 semesters), and after the course transformation (AT), from Spring 2018 through Fall 2019 (4 semesters). In each of these semesters, we collected responses at the start of the course (pre-test) and at the end of the course (post-test), and matched each pre response to a post response based on students' names and ID numbers. Students received course credit (1-2% of their final grade) for completing the surveys. A total of 1483 students in the BT semesters and a total of 1764 students in the AT semesters completed both pre and post surveys, and formed the data set analyzed here. Self-reported demographic information from these students is shown in Table I.

TABLE I. Self-reported gender, race/ethnicity, and major of participating students in E-CLASS surveys before (BT) and after (AT) course transformation.

	BT	AT
Females	25.8%	29.6%
Males	71.9%	68.0%
Gender non-conforming	1.4%	1.1%
Undeclared	0.9%	1.2%
American Indian or Alaska native	1.1%	0.8%
Asian American	16.5%	15.0%
Black or African American	1.6%	1.3%
Hispanic/Latino	8.9%	8.1%
Native Hawaiian or other Pacific Islander	0.7%	0.5%
White	69.5%	73.0%
Other race/ethnicity	3.9%	2.6%
Physics	7.6%	8.9%
Engineering	54.8%	52.4%
Other STEM	33.3%	34.6%
Other Disciplines	4.3%	4.0%

Analysis Methods. Students rank E-CLASS items on a 5-point Likert scale from “strongly disagree” to “strongly agree.” Responses to each item are scored based on how closely they align with expert-like responses. Following other previous studies on E-CLASS data [14, 18], we collapse “strongly (dis)agree” and “(dis)agree” into a single category, resulting in a 3-point scale. A response is then assigned a score of +1 if it aligns with the expert-like response, 0 for the neutral response, and -1 if it goes against the expert-like response. In addition to analyzing responses item-by-item, we also calculate each student's total score on the E-CLASS by summing their scores on each of the 30 items, resulting in an overall score with a possible range from -30 to 30.

To compare distributions of total scores, both pre to post and BT to AT, we use the nonparametric Mann-Whitney U test [19], with statistical significance determined at the 95% confidence level. To get a better sense of the differences be-

tween BT and AT semesters, we compare post scores while controlling for pre scores, as other work has shown significant correlation between a student's pre and post score on the E-CLASS [15]. We control for pre score using an analysis of covariance (ANCOVA) [20], with post score as the dependent variable, BT/AT as the independent variable, and pre score as a covariate. ANCOVA requires certain assumptions to be met, namely independence, homogeneity of variance, normality, linearity, independence of the covariate, and independent variable and homogeneity of regression slopes. Our data set meets each of these assumptions except for normality. However, ANCOVA is robust against violation of the normality assumption especially if the sample size is large [20], as is the case here. As above, we determine statistical significance at the 95% confidence level.

We perform an ANCOVA both on total scores and on scores to individual items. When comparing results item-by-item, we adjust results using the Bonferroni correction [21] to account for the problem of multiple comparisons. Lastly, as a measure of practical significance, we compute effect sizes using Cohen's d [22] for statistically significant differences. We take results with $d > 0.1$ as practically significant.

IV. RESULTS AND DISCUSSION

Overall E-CLASS Score. We calculated the overall E-CLASS score to investigate the difference in students' views and expectations about experimental physics in the BT and the AT courses. The overall average BT pre-test score is 18.6 ± 0.3 and the post-test score is 17.9 ± 0.3 (all reported uncertainties are given as 95% confidence intervals). The difference between the scores is statistically significant ($p < 0.001$) with Cohen's d of 0.13. For the AT courses, the pre-test score is 19.0 ± 0.3 and post-test score is 18.9 ± 0.2 ($p = 0.34$). Unlike the case for the BT courses, we do not find a statistically significant difference between the pre- and post-tests for the AT averages. The pre-test and post-test BT and AT scores are shown in Fig. 1. It is encouraging to note that the course transformation does not show a decrease unlike the untransformed course. Typically, students score more novice-like on attitudinal surveys and no change from pre to post is considered a relatively positive outcome. Of course, more expert-like views are the ultimate goal.

As the demographics are the same between the BT and AT classes, we expect the distributions of the pre-test scores to be the same. To check this, we compare the BT pre-test overall average score with that of the AT. We find that these scores are similar with a statistically significant difference between them ($p = 0.02 < 0.05$), but with a negligible Cohen's d of $= 0.07$. This suggest students hold similar views about experimental physics before instruction for both BT and AT courses, but that we should account for small differences in the pre-test scores in our analysis.

To take pre-test scores into account when evaluating the impact of course transformation, we perform an ANCOVA to compare the post-test average scores while controlling for the pre-test scores. This analysis is done on the matched data

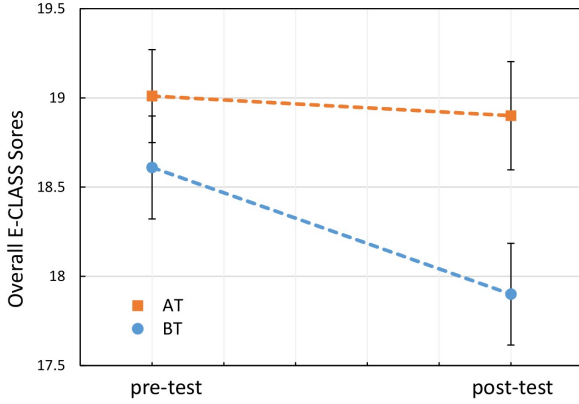


FIG. 1. Comparison of the average overall E-CLASS scores between the BT (1483 students) and AT (1764 students) offerings of the course. The statements associated with each Item number can be found in Ref [9].

set ($N_{BT} = 1483$ and $N_{AT} = 1764$). The results of the ANCOVA suggest a statistically significant effect of the covariate, pre-test scores, on the dependent variable, post-test scores ($p \ll 0.01$). There is also a statistically significant effect for the course type ($p \ll 0.01$). Controlling for the pre-test, the BT and AT adjusted average post-test scores are 18.1 ± 0.3 and 18.8 ± 0.3 , respectively. The difference between the average scores is 0.7 (down from 1.00 before controlling for the pre-test). The increased post-test overall score in AT courses is an encouraging sign for the transformation's effect. However, more insight can be gained by looking at data item-by-item.

Individual E-CLASS items. Previous studies [17] have cautioned looking at only overall scores, as not all E-CLASS items may be relevant to the course learning outcomes. For our study, only nine of the E-CLASS items have been previously identified as directly related to the BT course learning goals (citation redacted) Thus, we perform an ANCOVA with the pre-test score as a covariate for each individual E-CLASS item. The results of this analysis are shown in Fig. 2. In this figure, we sort the E-CLASS items on the horizontal axis in ascending order of post-test BT scores. It is clear that students respond more expert-like on many items in the AT courses as compared to the BT courses. It should be noted that items with high BT post-test scores (> 0.9) are in almost perfect agreement with expert-like views, which leaves little room for further improvement by the transformation.

Eleven items, indicated by asterisks in Fig. 2, show a statistically significant difference between the BT and AT post-test scores. These items along with their effect sizes, d , are listed in Table II. Out of these, five are directly relevant to the learning goals of the course. We note that Ref. [10] found, with a smaller data set, only three E-CLASS items relevant to the learning goals that showed a statistically significant shift from BT to AT.

To obtain a more complete understanding of the impact of

the transformation, we consider these 11 items in Table II and discuss possible reasons for the positive outcome based on the course structure and activities for a subset of these items.

- Item 16: For this item, a possible explanation for the positive results is that in the BT course all of the lab activities were to measure well-known quantities (e.g., index of refraction of Lucite), while in the AT course, the goal was never to measure a known value. For example, for the Snell's Law lab, students are now asked to determine the unknown concentration of sugar water. Additionally, different groups are given different concentrations, so the goal is not just to determine the concentration, but also to determine (using measurements and associated uncertainties) which other groups have the same concentration.
- Item 5: Statistical uncertainty analysis is emphasized throughout the AT course. Since confirmation of previously known results is not part of the AT course, students must use the uncertainty in their measured values to make scientific arguments or predictions. For example, in a lab where a projectile is fired through layers of tissue, students must use the uncertainty of measurements of the energy it takes to break through one tissue to determine the maximum number of tissues the projectile can break through.
- Item 23: Making predictions was an explicit step in most labs and was called out in lab guide headings. For example, one lab has students measure the exit velocity of a ball from the launcher and then use it to predict the the mean landing spot when the launcher is placed at an angle to the floor. They predict not just the average landing position, but also the range corresponding to the uncertainty in the launch velocity. This also reinforces the ideas associated with Item 5.
- Items 17/19: Due to the large range of instructional skill and motivation of the teaching assistants (TAs) in the lab, we designed the course to be successful even in the case of a non-expert TA. To accomplish this, there are many "check-in" points throughout the lab where a lab group of two students is asked to talk to another lab group about procedures, data analysis, interpretations, or results. Thus, we hoped to build a community in the lab where students felt comfortable asking other students questions. Additionally, for several labs, students are asked to enter their results into a common spreadsheet. Then, near the end of the lab time, students reason about conclusions that can be made from the data set as a whole rather than just their individual results. We suggest this structure may have impacted students' ideas around dealing with challenges in the lab and their ideas about group work.

V. CONCLUSIONS AND FUTURE WORK

The introductory lab course at CU has been transformed using a research-based approach, which included development of consensus learning goals and repeated assessment

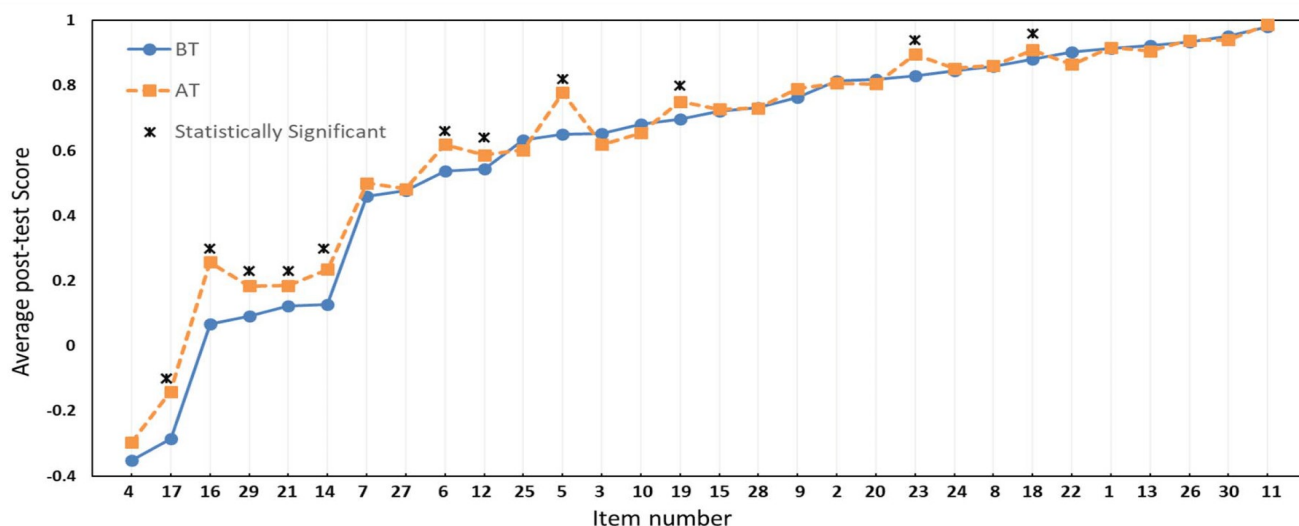


FIG. 2. Post-test average scores for each E-CLASS question for the BT and AT offerings of the course. E-CLASS items on the horizontal axis are ordered in ascending order of post-test BT scores. The points marked with an * indicate a statistically significant difference between the BT and AT versions of the course.

TABLE II. E-CLASS items that show a statistically significant difference between BT and AT scores. Items in bold are related to the courses learning goals. These items are ordered according to descending order of their effect sizes.

Number	Items	<i>p</i> -value	Effect size <i>d</i>
16	The primary purpose of doing physics experiments is to confirm previously known results.	< 0.001	0.24
5	Calculating uncertainties usually helps me understand my results better.	< 0.001	0.23
23	When I am doing an experiment, I try to make predictions to see if my results are reasonable.	< 0.001	0.16
17	When I encounter difficulties in the lab, my first step is to ask an expert, like the instructor.	0.002	0.15
14	When doing an experiment, I usually think up my own questions to investigate.	< 0.001	0.15
6	Scientific journal articles are helpful for answering my own questions and designing experiments.	< 0.001	0.15
21	I am usually able to complete an experiment without understanding the equations and physics idea that describe the system I am investigating.	0.001	0.13
19	Working in a group is an important part of doing physics experiments.	0.003	0.11
29	If I do not have directions for analyzing data, I am not sure how to choose an appropriate analysis method.	0.007	0.11
18	Communicating scientific results to peers is a valuable part of doing physics experiments.	0.013	0.10
12	I do not need to understand how the measurement tools and sensors work in order to carry out an experiment.	0.025	0.06

of the achievement of those goals. Progress towards meeting one of these goals, students' development of expert-like epistemology, was measured using a research-based assessment, E-CLASS. Using the data from E-CLASS, we have analyzed the impact of the transformation using a large data set of student responses both before and after the transformation. We see that transformation efforts lead to students' overall views not being negatively impacted by the course. Additionally, five of the nine E-CLASS items relevant to the course's learning goals show a statistically significant more expert-like shift for the transformed course. These improvements can be linked to course activities and structure.

Future work with these data will include analysis that looks at other variables such as gender, student interest in physics,

student major, and race/ethnicity. These factors could show interesting underlying dynamics in relation to the course transformation, which could motivate changes to the lab activities and course structure to better serve a diverse student population.

ACKNOWLEDGMENTS

NS would like to thank HJL for her kind hospitality at CU where this work has been done and would also like to thank Sultan Qaboos University for granting a sabbatical leave support at CU. This work was supported by NSF under Grant number PHYS-1734006.

-
- [1] J. Kozminski, H. J. Lewandowski, N. Beverly, S. Lindaas, D. Deardorff, A. Reagan, R. Dietz, R. Tagg, M. EblenZayas, J. Williams, H. R., and B. M. Zwickl, AAPT, College Park, MD (2014).
 - [2] N. G. Holmes and E. M. Smith, *The Physics Teacher* **57**, 296 (2019).
 - [3] A. Carter, in *2018 Conference on Laboratory Instruction Beyond the First Year of College*, BFY Conference (Baltimore, MD, 2018).
 - [4] in *Physics Education Research Conference 2016*, PER Conference (Sacramento, CA, 2016) pp. 104–107.
 - [5] D. Doucette, R. Clark, and C. Singh, *Physics Education Research Conference Proceedings* (2018).
 - [6] H. Greener, A. Vigder, S. Beck, and R. Beck-Barkai, *Physics Today* **71**, 12 (2018).
 - [7] E. M. Smith, M. M. Stein, C. Walsh, and N. G. Holmes, *Phys. Rev. X* **10**, 011029 (2020).
 - [8] H. J. Lewandowski, B. Pollard, and C. G. West, in *Physics Education Research Conference 2019*, PER Conference (Provo, UT, 2019).
 - [9] B. M. Zwickl, T. Hirokawa, N. Finkelstein, and H. J. Lewandowski, *Phys. Rev. ST Phys. Educ. Res.* **10**, 010120 (2014).
 - [10] B. Pollard and H. J. Lewandowski, in *2018 Physics Education Research Conference Proceedings* (2018).
 - [11] The process of transforming an advanced lab course: Goals, curriculum, and assessments, *American Journal of Physics* **81**, 63 (2013), <https://doi.org/10.1119/1.4768890>.
 - [12] T. S. Volkwyn, S. Allie, A. Buffler, and F. Lubben, *Phys. Rev. ST Phys. Educ. Res.* **4**, 010108 (2008).
 - [13] H. Lewandowski, D. R. Bolton, and B. B. Pollard, *Physics Education Research Conference Proceedings* (2018).
 - [14] B. R. Wilcox and H. J. Lewandowski, *Am. J. Phys.* **86**, 212 (2018).
 - [15] B. R. Wilcox and H. J. Lewandowski, *Phys. Rev. Phys. Educ. Res.* **12**, 010123 (2016).
 - [16] B. R. Wilcox and H. J. Lewandowski, *Phys. Rev. Phys. Educ. Res.* **13**, 020110 (2017).
 - [17] B. R. Wilcox and H. J. Lewandowski, *Phys. Rev. Phys. Educ. Res.* **12**, 020130 (2016).
 - [18] R. Henderson, K. Funkhouser, and M. D. Caballero, *Physics Education Research Conference Proceedings* (2019).
 - [19] H. B. Mann and D. R. Whitney, *Ann Math Stat* **18**, 50 (1947).
 - [20] D. L. Hahs-Vaughn and R. G. Lomax, *An Introduction to Statistical Concepts* (Routledge, New York, 2020).
 - [21] O. J. Dunn, *Journal of the American Statistical Association* **56**, 52 (1961).
 - [22] J. Cohen, *Statistical power analysis for the behavioral sciences* (L. Erlbaum Associates, 1988).