

# Impact of instructional approach on students' epistemologies about experimental physics

Bethany R. Wilcox<sup>1</sup> and H. J. Lewandowski<sup>1,2</sup>

<sup>1</sup>*Department of Physics, University of Colorado, 390 UCB, Boulder, CO 80309*

<sup>2</sup>*JILA, National Institute of Standards and Technology and University of Colorado, Boulder, CO 80309*

Student learning in undergraduate physics laboratory courses has garnered increased attention within the PER community. Considerable work has been done to develop curricular materials and pedagogical techniques designed to enhance student learning within laboratory learning environments. Examples of these transformation efforts include the Investigative Science Learning Environment (ISLE), Modeling Instruction, and integrated lab/lecture environments (e.g., studio physics). In addition to improving students' understanding of the physics content, lab courses often have an implicit or explicit goal of increasing students' understanding and appreciation of the nature of experimental physics. We examine students' responses to a laboratory-focused epistemological assessment – the Colorado Learning Attitudes about Science Survey for Experimental Physics (E-CLASS) – to explore whether courses using transformed curricula or pedagogy show more expert-like student epistemologies relative to courses using traditional guided labs, as well as how this trend varies based on student major or gender. Data for this study are drawn from an existing data set of responses to the E-CLASS from multiple courses and institutions.

## I. INTRODUCTION

Introductory laboratory courses represent a unique and important component of the physics curriculum [1] and have been specifically highlighted as critical pieces of the undergraduate curriculum by multiple professional groups [2–4]. For many students, these courses provide their first exposure to physics as an experimental science. As such, these courses can be gateways in terms of both recruiting and retaining students in the physics major. Moreover, for students who do not continue on to major in physics or another science, introductory laboratory courses often represent one of the only experiences these student will get with the process of scientific experimentation. Gaining a general understanding of the process and nature of scientific experimentation is an important part of creating citizens who are informed consumers of scientific information.

For these reasons, introductory physics lab courses often have explicit and/or implicit goals that go beyond just conveying physics content to developing expert-like epistemologies with respect to the nature of experimental physics [1, 2]. However, the traditional guided lab approaches to instruction commonly used in physics lab courses have been heavily critiqued as being cookbook and inauthentic [5, 6]. Inauthentic lab activities can have the unintended side-effect of encouraging epistemologies and expectations about the nature of experimental physics that are inconsistent with those of experts. Several groups within the PER community have worked to address this issue by designing laboratory learning environments that allow students to engage more authentically in the process of experimental physics. Examples of these environments include the Investigative Science Learning Environment (ISLE) [7], Modeling Instruction [8], and integrated lab/lecture environments such as stu-

dio physics [9] or SCALE-UP (Student-Centered Activities for Large Enrollment University Physics) [10]. All of these instructional approaches were either designed with improving students epistemologies about the nature of science as an explicit goal [7, 11], or have documented improved shifts in students epistemologies as measured by the CLASS (Colorado Learning Attitudes about Science Survey [12]) or MPEX (Maryland Physics Expectation Survey [13]) [14, 15].

Here, we investigate the potential impact of these transformed approaches to laboratory instruction on students' epistemologies and expectations about the nature of *experimental* physics. To do this, we use the Colorado Learning Attitudes about Science Survey for Experimental Physics (E-CLASS) [16, 17]. E-CLASS is a 30 item, Likert-style survey in which students are presented with a prompt (e.g., “The primary purpose of doing physics experiments is to confirm previously known results.”) and asked to rate their level of agreement both from their personal perspective when doing experiment in class and that of a hypothetical experimental physicist. The E-CLASS was developed in conjunction with laboratory course transformation efforts at the University of Colorado Boulder (CU) [18]. E-CLASS was validated through student interviews and expert review [16], and was tested for statistical validity and reliability using responses from students at multiple institutions and at multiple course levels [17]. This work is part of ongoing analysis of a growing, national data set of student responses to the E-CLASS.

## II. DATA SOURCES & METHODS

Data for this study were pulled from an existing data set of student responses to the E-CLASS collected us-

ing the E-CLASS online administration system [19] between 2013 and 2016. In order to use this system, instructors completed a Course Information Survey (CIS) at the beginning of the course. The CIS asked instructors to report logistical and demographic information about the course including instructional approaches used and course level (i.e., first year or beyond first year). To explore how students' E-CLASS responses might vary across demographic lines, the system also asked students to provide demographic information about themselves including their major and gender. While students were offered 15 distinct major options, we focus here on students' major as simply 'physics' or 'non-physics,' where physics includes both engineering physics and physics majors. The primary motivation for this collapsing student majors into these binary categories was to preserve statistical power. To account for the complex and non-binary nature of gender [20], students were provided three response options when reporting their gender: woman, man, and other. Roughly 2% ( $N = 87$ ) of the students selected the 'other' option or left the question blank; thus, to preserve statistical power, the analysis with respect to gender will include only those students who selected 'man' or 'woman' when reporting their gender.

As the transformed instructional approaches described above have been designed primarily for introductory courses, we limited our analysis to the subset of E-CLASS data drawn from first-year labs covering introductory physics content. Only students with matched pre- and post-instruction responses were included in the analysis. The final, matched data set included  $N = 3601$  student responses from 60 separate instances of 44 distinct algebra- or calculus-based first-year courses spanning 29 institutions. Based on the instructors' responses on the CIS, each of these courses was classified as either Transformed or Traditional. Courses in the Transformed category were those whose instructors reported using one or more of the transformed instructional approaches described above – ISLE, Modeling, Studio Physics, or SCALE-UP. Traditional courses were those whose instructors reported using only a traditional guided lab approach to instruction. There were 20 Transformed and 40 Traditional courses in the data set; Table I shows the

TABLE I. Breakdown of students by major and gender in the matched first-year data set. 'Physics' includes both physics and engineering physics majors, and Non-physics includes other engineering, science, math, non-science, and undeclared. A small number of students did not provide their major or gender; thus, the percentages will not sum to 100%.

	N	Major		Gender	
		Physics	Non-Physics	Woman	Man
Transformed	510	7%	91%	42%	56%
Traditional	3091	8%	90%	42%	55%

breakdown of students by major and gender across these courses.

For scoring purposes, students' responses to each 5-point Likert item were condensed into a standardized, 3-point scale in which the responses '(dis)agree' and 'strongly (dis)agree' were collapsed into a single category [12]. Students' responses to individual items were given a numerical score based on consistency with the accepted, expert response: favorable (+1), neutral (+0), or unfavorable (−1). A student's overall score on the assessment is given by the sum of their scores on each of the 30 E-CLASS items resulting in a possible range of scores of  $[-30, 30]$ . For more information on the scoring of the E-CLASS see Ref. [17]. Because the distribution of scores on the E-CLASS is typically skewed towards positive scores, the following sections report statistical significance based on the non-parametric Mann-Whitney U test [21] unless otherwise stated. Where differences between means are statistically significant, we also report Cohen's  $d$  [22] as a measure of effect size.

### III. RESULTS

The average overall pre- and post-instruction E-CLASS scores for Transformed and Traditional courses are given in Table II. Students in Traditional courses had a statistically significant ( $p \ll 0.01$ ,  $d = -0.3$ ), negative shift from pre- to post-instruction. This negative shift over the course of one semester or quarter is consistent with historical trends for the E-CLASS [17]. Alternatively, students in Transformed courses did not shift significantly from pre- to post-instruction ( $p = 0.2$ ). As shown in Table II, the difference between the pre-instruction scores for Traditional and Transformed was not statistically significant; however, the difference in post-instruction scores was significant with a moderate effect size.

We can also examine the difference between Transformed and Traditional courses on an item-by-item scale. Fig. 1 shows the difference between the average scores of students in Transformed and Traditional course for each of the 30 items on the pre- and post-instruction E-CLASS. Consistent with the results from the overall score, Fig. 1 shows that only two items show a statistically significant difference between the pre-instruction

TABLE II. Raw pre- and post-instruction E-CLASS scores (points) for Transformed and Traditional courses. Here, 'Significance' refers to the statistical significance of the difference between the averages for students in Transformed and Traditional courses (Mann-Whitney U test).

	Traditional	Transformed	Significance	Effect size
Pre	16.4	17.1	$p = 0.07$	
Post	14.4	17.3	$p \ll 0.01$	$d = 0.4$

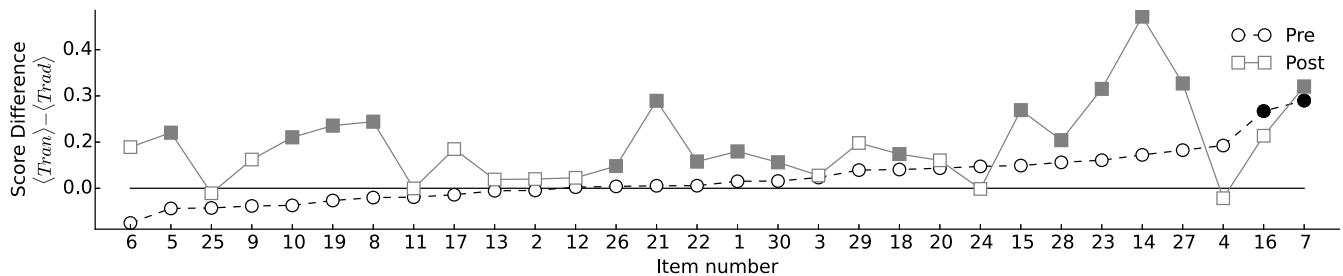


FIG. 1. Sorted plot of the difference between the average scores (points) of students in Transformed and Traditional courses for each item of the E-CLASS. Zero difference is marked by the solid horizontal line. Filled markers indicate points for which the difference between the students in Traditional and Transformed courses is statistically significant (Mann-Whitney U [21] and Holm-Bonferroni [23] corrected  $p < 0.05$ ). See Ref. [24] for full list of item prompts.

scores for students when split by instructional approach. The magnitude of the difference in pre-instruction scores was small for both items (Cohen’s  $d = 0.2$ ). Alternatively, 16 of the 30 E-CLASS items show Transformed courses scoring significantly better at the end of the course (Mann-Whitney U [21] and Holm-Bonferroni [23] corrected  $p < 0.05$ ). The magnitude of the difference between post-instruction scores was moderate (Cohen’s  $d > 0.3$ ) for only two of these items (Items 14 – “When doing an experiment, I usually think up my own questions to investigate” and item 23 – “When I am doing an experiment, I try to make predictions to see if my results are reasonable”) and small for all other items ( $d < 0.3$ ). Students in Traditional classes did not have significantly higher post-instruction means on any items.

While comparison of the E-CLASS scores between Traditional and Transformed courses preliminarily indicates that Transformed courses achieved better gains overall and on the majority of items, previous work has shown that other factors, such as student major and gender, can also impact the comparison of overall E-CLASS means for certain subpopulations [25]. To more clearly disentangle the relationship between instructional approach and post-instruction E-CLASS scores from that of other factors (e.g., pre-score, major), we used an analysis of covariance (ANCOVA) [26]. ANCOVA is a statistical method for comparing the difference between population means while accounting for the variance associated with other variables. Only students for whom we have matched E-CLASS scores along with both major and gender data were included in the ANCOVA ( $N = 3481$ ).

We first ran a 4-way ANCOVA to compare post-instruction E-CLASS means for Traditional and Transformed courses while controlling for student major and gender, as well as pre-instruction score as a covariate. To determine how these variables might also be related to one another, we included in the ANCOVA all possible interaction terms. The 4-way ANCOVA revealed a statistically significant interaction between gender and instructional approach (F-test,  $p = 0.03$ ). The presence of this interaction suggests that the relationship between instructional approach and post-instruction mean is dif-

ferent for men and women.

To account for this interaction term, we split the data by gender and ran 3-way ANCOVAs for men and women separately. Student major was a significant predictor of post-instruction scores for both men and women (F-test,  $p = 0.01$ ). Instructional approach was also a significant predictor for both men and women (F-test,  $p \ll 0.01$ ), with a higher mean for students in Transformed courses in both cases. However, the increase in post-instruction mean in Transformed courses was higher for women (see Table III) and the difference was significant (as indicated by the statistical significance of the interaction between gender and instructional approach). This finding supports the idea that both men and women benefited from Transformed instructional approaches, and this benefit was larger for women than men.

Similarly, to determine the significance of gender as a predictor of post-instruction E-CLASS scores, we split the data by instructional approach and ran separate 3-way ANCOVAs for Transformed and Traditional courses separately. For students in Traditional courses, adjusted post-instruction means for men were significantly higher than the adjusted means for women (F-test,  $p < 0.01$ ); however, for students in Transformed courses, post-instruction means for men and women were the same (F-test,  $p = 0.4$ ). Thus, gender was a significant predictor of post-instruction E-CLASS means in Traditional courses only.

TABLE III. Adjusted post-instruction E-CLASS means (points) from the 3-way ANCOVAs run on data from men and women separately. Adjusted means were calculated while controlling for the variance associated with pre-instruction score and student major. ‘Difference’ represents the mean for students in Transformed courses minus the mean for students in Traditional courses.

	Traditional	Transformed	Difference
Women	12.8	16.0	3.2
Men	15.8	17.6	1.8

#### IV. CONCLUSIONS & LIMITATION

We analyzed student responses to a laboratory focused epistemologies assessment – the E-CLASS – to examine the relationship between the use of transformed instructional approaches and students’ performance. Data were drawn from a large-scale, national data set of students taking a first-year physics laboratory course. We found that students in courses taught using transformed pedagogy had better post-instruction E-CLASS scores than students in courses taught using only traditional guide lab approaches to instruction. This result was supported by the results of an ANCOVA, which examined the difference between post-instruction means while adjusting for the variance associated with pre-instruction scores, student major, and student gender. The ANCOVA analysis showed that, when accounting for the variance associated with pre-instruction score and major, students in courses using transformed approaches to instruction had higher post-instruction means than those in courses using traditional approaches. Moreover, the ANCOVA showed that this effect was larger for women than men suggesting that transformed pedagogies may have had a differentially positive effect on women. Additionally, while gender was a significant predictor for Transformed courses, the differentially positive impact for women in Transformed courses made it so that gender was not a significant predictor for students in these courses.

There are several limitations to the results reported

here. This study was limited to first-year (introductory) laboratory courses; additional data collection and analysis would be necessary to determine if these results hold in beyond-first-year courses. Additionally, in order to preserve statistical power, all transformed pedagogies considered here (ISLE, Modeling Instruction, studio physics, and SCALE-UP) were aggregated together as a single group. Thus, while it may be that one approach is more or less effective at promoting expert-like epistemologies and expectations, the current analysis cannot address this dynamic. As data collection with the E-CLASS centralized administration system continues, comparisons between different transformed pedagogies may become possible. Finally, the purely quantitative analysis reported here cannot speak to how these transformed instructional approaches may have improved students’ epistemologies relative to traditional approaches, nor can it determine why they differentially benefited women. Determining the mechanism underlying the findings reported here will likely require additional research with a significant qualitative component.

#### ACKNOWLEDGMENTS

This work was funded by the NSF-IUSE Grant DUE-1432204 and NSF Grant PHY-1125844. Special thank you to the members of PER@C for all their feedback.

- 
- [1] R. Trumper, *Science & Education* **12**, 645 (2003).
  - [2] AAPT Committee on Laboratories, “AAPT Recommendations for the Undergraduate Physics Laboratory Curriculum,” (2015).
  - [3] S. Olson and D. G. Riordan, Executive Office of the President (2012).
  - [4] NRC Committee on Undergraduate Biology Education, *BIO2010: Transforming undergraduate education for future research biologists* (National Academies Press (US), 2003).
  - [5] C. Wieman, *The Physics Teacher* **53** (2015).
  - [6] NRC, *Adapting to a Changing World—Challenges and Opportunities in Undergraduate Physics Education* (The National Academies Press, 2013).
  - [7] E. Etkina and A. V. Heuvelen, in *2001 PERC Proceedings* (Rochester, New York, 2001).
  - [8] M. Wells, D. Hestenes, and G. Swackhamer, *AJP* **63**, 606 (1995).
  - [9] J. M. Wilson, *The Physics Teacher* **32**, 518 (1994).
  - [10] R. J. Beichner, J. M. Saul, R. J. Allain, D. L. Deardorff, and D. S. Abbott, 2000 Proceedings of the ASEE (2000).
  - [11] R. J. Beichner, J. M. Saul, D. S. Abbott, J. J. Morse, D. Deardorff, R. J. Allain, S. W. Bonham, M. H. Dancy, and J. S. Risley, *Research-based reform of university physics* **1**, 2 (2007).
  - [12] W. K. Adams, K. K. Perkins, N. S. Podolefsky, M. Dubson, N. D. Finkelstein, and C. E. Wieman, *PRST-PER* **2**, 010101 (2006).
  - [13] E. F. Redish, J. M. Saul, and R. N. Steinberg, *American Journal of Physics* **66**, 212 (1998).
  - [14] E. Brewster, L. Kramer, and G. O’Brien, *PRST-PER* **5**, 013102 (2009).
  - [15] P. B. Kohl and H. Vincent Kuo, *American Journal of Physics* **80**, 832 (2012).
  - [16] B. M. Zwickl, T. Hirokawa, N. Finkelstein, and H. Lewandowski, *PRST-PER* **10**, 010120 (2014).
  - [17] B. R. Wilcox and H. J. Lewandowski, *PR-PER* **12**, 010123 (2016).
  - [18] B. M. Zwickl, N. Finkelstein, and H. Lewandowski, *AJP* **81**, 63 (2013).
  - [19] B. R. Wilcox, B. M. Zwickl, R. D. Hobbs, J. M. Aiken, N. M. Welch, and H. J. Lewandowski, *ArXiv e-prints* (2016), arXiv:1601.07896 [physics.ed-ph].
  - [20] A. L. Traxler, X. C. Cid, J. Blue, and R. Barthelemy, *ArXiv e-prints* (2015), arXiv:1507.05107 [physics.ed-ph].
  - [21] H. B. Mann and D. R. Whitney, *The annals of mathematical statistics*, 50 (1947).
  - [22] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences* (L. Erlbaum Associates, 1988).
  - [23] S. Holm, *Scandinavian journal of statistics*, 65 (1979).
  - [24] [tinyurl.com/ECLASS-physics](http://tinyurl.com/ECLASS-physics), (2015).
  - [25] B. R. Wilcox and H. J. Lewandowski, *ArXiv e-prints* (2016), arXiv:1606.02629 [physics.ed-ph].
  - [26] A. R. Wildt and O. Ahtola, *Analysis of covariance*, Vol. 12 (Sage, 1978).