Correlation of open-ended activities in laboratory courses with students' views of experimental physics

Qiaoyi Liu[1,2] and H. J. Lewandowski[1,2]

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

Improving students' views of experimental physics is often an important goal of undergraduate physics laboratory courses. However, traditional lab courses typically include highly guided activities that often do not require or encourage students to engage in the authentic process of experimental physics. Alternatively, open-ended activities in lab courses can provide students with a more authentic learning experience. Here, we investigate the impact of open-ended activities in lab courses on students' views of experimental physics, as measured by the Colorado Learning Attitudes about Science Survey for Experimental Physics (E-CLASS), and concluded that the inclusion of some open-ended activities is associated with more expert-like post-instruction responses relative to the courses that include only traditional guided activities, and the effect is larger for students with low pre-instruction scores. We also found that the number of weeks spent on open-ended activities is not associated with pre-to-post instruction gain in E-CLASS scores.

2024 PERC Proceedings edited by Ryan, Pawl, and Zwolak; Peer-reviewed, doi.org/10.1119/perc.2024.pr.Liu

I. INTRODUCTION

Physics laboratory courses are considered to be an important component of the undergraduate curriculum [1]. These courses can provide students with valuable opportunities to engage in authentic scientific practices, develop practical lab skills, and engage collaboratively with other students.

A large portion of undergraduate physics lab courses are currently taught using only traditional guided lab activities. These highly structured labs have been critiqued as being inauthentic to the process of experimental physics [2, 3]. In response, members of the physics education research community developed several new pedagogical approaches specifically designed to allow students to engage in the process of experimental physics in a more authentic fashion. A key feature of these pedagogical approaches is the inclusion of open-ended lab activities, where students are provided with the opportunity for more agency in making decisions about the experiment. Examples of widely used pedagogical approaches that incorporate open-ended hands-on activities include the Investigative Science Learning Environment (ISLE) [4], Modeling Instruction [5], studio physics [6], Student-Centered Activities for Large Enrollment University Physics (SCALE-UP) [7], and Thinking Critically in Physics Labs [8]. Additionally, there are many lab courses that incorporate open-ended activities that have been developed at a single institution. Although previous literature presents how effective some of these pedagogical approaches are in terms of achieving their respective learning goals, there is still much left unstudied when it comes to quantifying how open-ended activities in lab courses can improve students' views of experimental physics, which is a common learning goal for most lab courses [1, 9].

Wilcox *et al*. [10] have previously explored the impact of open-ended activities in undergraduate lab courses on students' views of experimental physics, as measured by a laboratory-focused assessment known as the Colorado Learning Attitudes about Science Survey for Experimental Physics (E-CLASS) [11–13]. E-CLASS is a research-based assessment that measures students' epistemologies and expectations about experimental physics, as well as student affect and confidence when performing physics experiments. It presents students with a total of 30 statements (for instance, "When doing an experiment, I try to understand how the experimental setup works.") and asks them to rate their level of agreement on a Likert scale both from their personal perspective when doing experiments in class and that of a hypothetical experimental physicist. It was validated through student interviews and expert review [14], and was tested for statistical validity and reliability using responses from students at multiple institutions and at multiple course levels [15].

Wilcox *et al*. [10] have shown that the inclusion of some open-ended activities in a lab course correlates with more expert-like responses after instruction as compared to courses that include only guided activities. However, the extent to which students' views on experimental physics can be improved by a certain amount of open-ended activities included in the lab course is not explored in this previous work, as it aggregated all courses that use any open-ended activities together in an effort to preserve statistical power. Understanding how much open-ended activities are needed to see these positive results could serve as a guide for lab course instructors to incorporate sufficient amount of open-ended activities that could significantly improve their students' views on experimental physics.

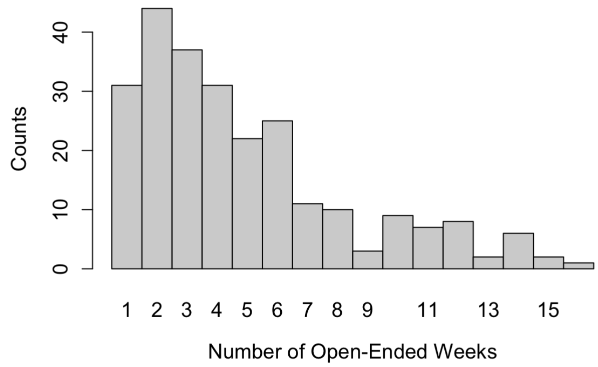

FIG. 1. Histogram of the number of weeks spent on open-ended activities for a total of 202 lab courses that include at least one week of open-ended activities.

Building on Wilcox *et al*. [10], our study utilizes a larger number of student responses to the E-CLASS survey, which allows us to take a more detailed look at how open-ended activities impact E-CLASS scores. Unlike in the previous study [10], we are able to probe how the *number of weeks* of open-ended activities correlates with E-CLASS scores. We aim to answer the following two research questions:

- RQ1: To what extent do open-ended activities in lab courses positively correlate with students' views of experimental physics as compared to courses with only guided labs? Does the correlation depend on their initial views before the course?
- RQ2: How many open-ended activities, as measured by the number of weeks in a lab course, are necessary to see an improvement in students' views of experimental physics?

II. METHOD AND DESIGN

A. Data Source

The data used for this study were collected from undergraduate physics lab courses between Spring 2015 and Fall 2023 using the E-CLASS centralized administration system [16]. Students completed the E-CLASS survey both pre- and post-instruction, typically in the first and last week of the course, respectively. During this period, we col-

TABLE I. Overall mean E-CLASS scores and their standard errors (SEs) on both the pre- and post-tests for students in courses using only guided activities and those using open-ended activities. Significance and Cohen's d, reported as 95% confidence interval (CI), describe the difference between students' scores in the guided-only and open-ended courses.

	Guided Only	**Some Open-Ended**	**Significance**	**Cohen's d (95% CI)**
Number of Students	9852	4104	-	-
Pre E-CLASS Score (Mean ± SE)	16.38 ± 0.07	16.94 ± 0.10	$p < 0.001$	[0.065, 0.103]
Post E-CLASS Score (Mean ± SE)	15.12 ± 0.07	17.20 ± 0.10	$p < 0.001$	[0.252, 0.290]

lected a total of 70017 pre-instruction responses and 55193 post-instruction responses. Only students for whom we had matched pre- and post-instruction responses were included in the analysis. The E-CLASS survey also includes a filtering question to eliminate responses from students who did not read the item prompts, so that any student who responded incorrectly to this filtering question was also removed from the analysis [15, 17]. The E-CLASS matched and filtered data set includes a total of 40129 matched student responses. To quantify the amount of lab activities included in the course, we asked instructors to report how many weeks of instruction were spent on "all guided lab activities" and how many weeks were spent on "all open-ended activities or projects", and 531 out of 1080 courses reported this information. We also removed courses that reported more than 17 weeks total of instruction, as those are likely year-long lab courses, or simply errors. This all resulted in 13956 matched student responses in 389 distinct courses.

Out of the 389 courses, 187 of them had only guided lab activities, while 202 of them included at least one week of open-ended activities. The distribution of the weeks spent on open-ended activities are shown in Fig. 1. While the number of weeks spent on open-ended activities varied significantly amongst the courses, the distribution overall is skewed to lower number of weeks with an overall mean of 4.88 weeks, indicating that most courses include only a relatively small amount of open-ended activities compared to typical 10 – 15 week terms.

B. Analysis Method

Response options for E-CLASS items are given on a 5-point Likert scale, from "strongly agree" to "strongly disagree". For scoring purposes, students' responses to each 5-point E-CLASS item were condensed into a standardized, 3-point scale in which the responses "(dis)agree" and "strongly (dis)agree" were collapsed into a single (dis)agree category. Students' responses to individual items were given a score based on consistency with the expert response: +1 for a response consistent with experts, 0 for neutral, and -1 for a response inconsistent with experts [15]. A student's overall E-CLASS score is then given by the sum of their scores on each of the 30 items resulting in a possible score range of [-30, 30].

To measure the impact of open-ended activities in lab courses as compared to the fully guided ones, we first treated the 389 courses dichotomously as either having open-ended activities, regardless of the number of weeks spent on those activities, or having only guided activities. We first examine the overall behavior by comparing the means of the E-CLASS scores pre- and post-instruction for students in courses using open-ended activities and those using only guided activities. As the distribution of scores on the E-CLASS is typically skewed towards positive scores [15], we report statistical significance based on the non-parametric Wilcoxon signed-rank test [18] unless otherwise stated. In the case where the differences between the means are statistically significant, we also report the Cohen's d [19] in terms of the 95% confidence interval (CI) as a measure of effect size and practical significance. To explore this effect in greater detail, we also utilize an analysis of covariance (ANCOVA) [20] in addition to examining students' raw pre- and post-instruction E-CLASS scores. ANCOVA is a statistical method for comparing the difference between population means after adjusting them to account for the variance associated with other variables. In this case, we want to determine whether the difference between the E-CLASS scores of students in courses using different types of lab activities, i.e. open-ended vs. guided only, remains statistically significant after accounting for differences in pre-instruction scores.

To measure the amount of open-ended activities that is correlated with a significant change in students' views of experimental physics, we constrain our analysis to a subset of the full data set where we include only the 202 courses with at least one week of open-ended activities. We chose the number of weeks spent on open-ended activities as the measure of the amount of open-ended activities included during the the course. The rationale behind choosing the number of weeks spent on open-ended activities as the variable of interest, as opposed to fraction of weeks on open-ended activities, is that the total number of weeks are not the same for all the courses. Since the number of weeks is a continuous variable, we utilized a nested set of linear regression models to control for students' pre-instruction scores.

III. ANALYSIS AND RESULTS

A. RQ1: Comparing fully guided courses to courses with open-ended activities

To explore general trends in the aggregate data, we first examine the differences in raw pre- and post-instruction E-

CLASS scores for students in courses using open-ended activities and those using only guided activities, as shown in Table I. According to Table I, students in courses using open-ended activities score significantly higher than those in courses using only guided activities both pre- and post-instruction ($p < 0.001$). While the difference is statistically significant both before and after instruction, the magnitude of this effect is much larger for the post-instruction scores. Moreover, students in courses using open-ended activities showed a small (95% CI [0.015, 0.060]), but statistically significant positive shift ($p < 0.001$) from before to after instruction, while those in courses using only guided activities showed a relatively larger (95% CI [-0.187, -0.158]), but statistically significant negative shift ($p < 0.001$). The comparison of E-CLASS scores between students in courses using different types of lab activities suggests that open-ended activities have a significant positive impact of students' views of experimental physics, as compared to the guided ones.

To explore the relationship between open-ended activities and post-instruction E-CLASS scores in a more detailed way, we performed a one-way ANCOVA to compare post-instruction E-CLASS scores for courses using guided and open-ended activities, while using the pre-instruction score as a covariate. In order for the results of an ANCOVA to be valid, the data must meet several assumptions [20], one of which is that the slope of the regression line between the dependent variable and covariate is the same for each category of the independent variable. To verify this assumption for the E-CLASS matched data, we initially employed the following model, which included an interaction term between pre-instruction score and the instruction type:

$$PostECLASSScore \sim PreECLASSScore + OpenEnded + PreECLASSScore \times OpenEnded \quad (1)$$

In this model, $PreECLASSScore$ and $PostECLASSScore$ correspond to students' E-CLASS scores pre- and post-instruction, while the categorical variable $OpenEnded$ corresponds to the instruction type, and is coded as 0 for students in courses using only guided activities and 1 for those in courses using open-ended activities. After fitting this model to the E-CLASS matched data, we found that there is a statistically significant interaction between pre-instruction score and the instruction type ($F = 23.23$, $p < 0.001$). The pre- vs. post-instruction E-CLASS scores for students in courses using guided and open-ended activities according to the ANCOVA results are plotted in Fig. 2, which further confirms that the slopes are unequal between the two instruction types.

Tests of the E-CLASS matched data showed that they satisfied all the ANCOVA assumptions, except homogeneity of regression slopes, as is indicated by the statistically significant interaction term. In light of this assumption violation, we subsequently ran the one-way ANCOVA model without the interaction term. After fitting this nested model to the E-CLASS matched data, we found that there is a statistically significant effect on the instruction type on the post-instruction

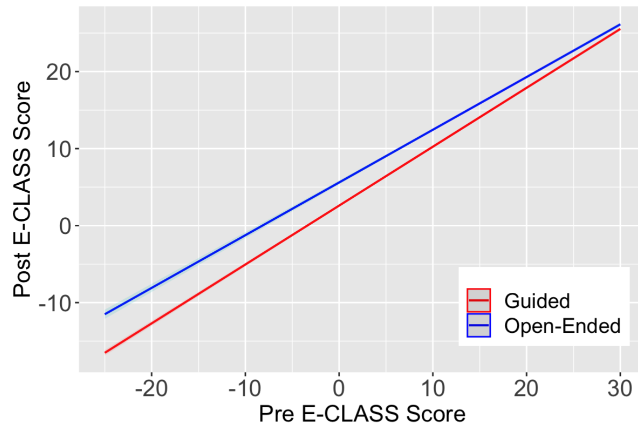

FIG. 2. Models for post-instruction vs. pre-instruction E-CLASS scores for students in courses using only guided activities (solid red line) and those using some open-ended activities (solid blue line).

score after controlling for pre-instruction score ($F = 231.19$, $p < 0.001$), although the effect size is fairly small (partial η^2 = 0.016). Since the homogeneity of slope assumption is violated, our results here should be interpreted as lower bounds for the partial η^2 instead of absolute values on the relationship between the instruction type and post-instruction E-CLASS score.

Overall, the results of the descriptive statistics and ANCOVA indicate that students in courses with open-ended activities all scored significantly higher than those in courses with only traditional guided activities after instruction while controlling for the pre-instruction score, but the shift is more significant for students with low pre-instruction scores, as shown in Fig. 2. This suggests that open-ended activities may have a greater benefit for students who started with less expert-like views of experimental physics.

B. RQ2: Comparing E-CLASS scores for varying number of weeks of open-ended activities

To provide a more detailed picture of how much open-ended activities can improve students' views of experimental physics, we investigated a subset of the E-CLASS matched data that includes only courses with at least one week spent on open-ended activities. We do this by fitting the following linear regression model to the data, which includes an interaction term between pre-instruction score and the number of weeks spent on open-ended activities:

$$PostECLASSScore = \beta_0 + \beta_{pre} PreECLASSScore + \beta_{open} OpenWeeks + \beta_{int} PreECLASSScore \times OpenWeeks \quad (2)$$

In this model, $OpenWeeks$ corresponds to the number

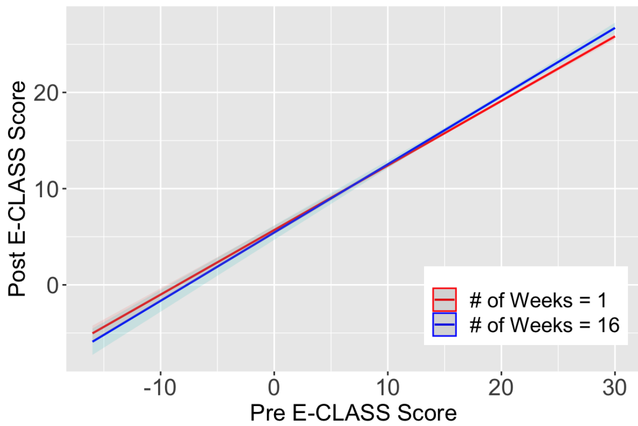

FIG. 3. Models for post-instruction vs. pre-instruction E-CLASS scores for students in courses with 1 week spent (solid red line) and those with 16 weeks spent (solid blue line) on open-ended activities.

of weeks spent on open-ended activities. After fitting this model to the subset of the E-CLASS matched data (excluding courses with $OpenWeeks = 0$), we found that there is no significant interaction between pre-instruction score and the number of weeks spent on open-ended activities ($\beta_{int} = 0.002 \pm 0.004$, $p = 0.502$). This indicates that the impact of open-ended activities is consistent for all students, those with both high and low pre-instruction scores.

Finding that the interaction term is not statistically significant, we removed it from the model, and reran the analysis without the interaction term. We then found that the number of weeks spent on open-ended activities does not significantly predict the post-instruction score after controlling for the pre-instruction score ($\beta_{pre} = 0.03 \pm 0.02$, $p = 0.228$), and the effect size is negligible (partial η^2 = 0.0003). The models for pre- vs. post-instruction E-CLASS scores for students in courses with the lowest (1) and highest (16) number of weeks spent on open-ended activities according to the linear regression results are plotted in Fig. 3. The overlap between the two fitted lines across all pre-instruction scores further confirms that the number of weeks spent on open-ended activities is not associated with pre-to-post-instruction gain in E-CLASS score.

Overall, the results of the linear regressions indicate that the amount of open-ended activities, as measured by the number of weeks, does not significantly impact students' views of experimental physics as long as that number is not zero.

IV. CONCLUSIONS AND FUTURE WORK

Based on the results of the quantitative analysis, we concluded that lab courses with open-ended activities can generally promote expert-like views of experimental physics better than those that include only traditional guided activities, especially for students with less expert-like initial views. We also found that the number of weeks of open-ended activities is not associated with pre-to-post-instruction gain in E-CLASS score, a distinct exception to the common trend observed in education research that more instruction results in better student performance [21, 22].

There are a few limitations to keep in mind when interpreting this work, which should be examined in future studies. First, the number of weeks spent on open-ended activities may not be the best proxy for the amount of open-ended activities included in a course. One method to extract this information is to evaluate the syllabi of these courses. Another is to leverage the additional data collected from the instructors regarding the frequency of decision-making, modeling, and communication-based activities in the course.

Furthermore, several factors are not controlled for in this study, but would be interesting to investigate separately. One is the selection effect on the instructor. Instructors who include at least one week of open-ended activities may value improving student belief more than those who do not. As such, even for courses with a low number of open-ended weeks, the guided portion may still have more open-ended-like pedagogical features, which needs to be tested in future studies. Another factor to be explored is the recency effect. The placement of open-ended activities in the course may vary, and this information is not well-represented solely by the number of open-ended activities. However, based on a brief evaluation of some of the course syllabi, the open-ended activities are generally aggregated toward the end of the course, in the form of a final project.

Finally, it is important to note that this study is exploratory in nature. Future controlled experiments are needed to establish the causal relationship between the amount of open-ended activities and improvement in student views of experimental physics.

Nevertheless, this study demonstrated that a large amount of open-ended activities may not be necessary to improve students' views of experimental physics. This has the potential to significantly lower the barrier for instructors to include open-ended activities in their lab courses in an effort to improve students' views of experimental physics.

ACKNOWLEDGMENTS

This research was primarily supported by NSF DUE-2142356 and PHY 2317149.

[1] R. Trumper, Sci. Educ. **12**, 645 (2003).
[2] C. Wieman, Phys. Teach. **53**, 349 (2015).
[3] N. G. Holmes, and C. Wieman, Phys. Today **71**, 38 (2018).
[4] E. Etkina, and A. Van Heuvelen, in *2001 PERC Proceedings, Rochester, NY, 2001*, edited by K. Cummings, S. Franklin, and J. Marx.
[5] M. Wells, D. Hestenes, and G. Swackhamer, Am. J. Phys. **63**, 606 (1995).
[6] J. M. Wilson, Phys. Today **32**, 518 (1994).
[7] R. J. Beichner, J. M. Sau, R. J. Allain, D. L. Deardorff, and D. S. Abbott, in *Proceedings of the Annual Meeting of the American Society for Engineering Education, Seattle, Washington, 2000*.
[8] R. E. Scherr, and N. G. Holmes, CourseSource **10** (2023).
[9] B. Zwickl, N. Finkelstein, and H. J. Lewandowski, Am. J. Phys. **81**, 63 (2013).
[10] B. R. Wilcox, and H. J. Lewandowski, Phys. Rev. Phys. Educ. Res. **12**, 020132 (2016).
[11] B. Zwickl, N. Finkelstein, and H. J. Lewandowski, AIP Conf. Proc. **1513**, 442 (2013).
[12] B. R. Wilcox, and H. J. Lewandowski, Phys. Rev. Phys. Educ. Res. **13**, 020110 (2017).
[13] B. R. Wilcox, and H. J. Lewandowski, Am. J. Phys. **86**, 212 (2018).
[14] B. Zwickl, T. Hirokawa, N. Finkelstein, and H. J. Lewandowski, Phys. Rev. ST Phys. Educ. Res. **10**, 010120 (2014).
[15] B. R. Wilcox, and H. J. Lewandowski, Phys. Rev. Phys. Educ. Res. **12**, 010123 (2016).
[16] B. R. Wilcox, B. M. Zwickl, R. D. Hobbs, J. M. Aiken, N. M. Welch, and H. J. Lewandowski, Phys. Rev. Phys. Educ. Res. **12**, 010139 (2016).
[17] J. M. Aiken, and H. J. Lewandowski, Phys. Rev. Phys. Educ. Res. **17**, 020144 (2021).
[18] F. Wilcoxon, Biometrics Bulletin. **1**, 80 (1945).
[19] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences* (L. Erlbaum Associates, Hillsdale, NJ, 1988).
[20] A. R. Wildt, and O. Ahtola, *Analysis of Covariance* (Sage, Thousand Oaks, CA, 1978), Vol. 12.
[21] D. Clark, and M. C. Linn, J. Learn. Sci. **12**, 451 (2003).
[22] S. C. Andersen, M. K. Humlum, and A. B. Nandrup, PNAS **113**, 7481 (2016).