PHYSICS EDUCATION RESEARCH SECTION

The Physics Education Research Section (PERS) publishes articles describing important results from the field of physics education research. Manuscripts should be submitted using the web-based system that can be accessed via the American Journal of Physics home page, http://ajp.dickinson.edu, and will be forwarded to the PERS editor for consideration.

A summary of research-based assessment of students' beliefs about the nature of experimental physics

Bethany R. Wilcox
Department of Physics, Colorado School of Mines, Golden, Colorado 80401

H. J. Lewandowski
Department of Physics, University of Colorado Boulder, Boulder, Colorado 80302 and JILA, National Institute of Standards and Technology, and University of Colorado, Boulder, Colorado 80309

(Received 26 May 2017; accepted 11 October 2017)

Within the undergraduate physics curriculum, students' primary exposure to experimental physics comes from laboratory courses. Thus, as experimentation is a core component of physics as a discipline, lab courses can be gateways in terms of both recruiting and retaining students within the physics major. Physics lab courses have a wide variety of explicit and/or implicit goals for lab courses, including helping students to develop expert-like beliefs about the nature and importance of experimental physics. To assess students' beliefs, attitudes, and expectations about the nature of experimental physics, there is currently one research-based assessment instrument available—the Colorado Learning Attitudes about Science Survey for Experimental Physics (E-CLASS). Since its development, the E-CLASS has been the subject of multiple research studies aimed at understanding and evaluating the effectiveness of various laboratory learning environments. This paper presents a description of the E-CLASS assessment and a summary of the research that has been done using E-CLASS data with a particular emphasis on the aspects of this work that are most relevant for instructors.
https://doi.org/10.1119/1.5009241

I. INTRODUCTION

Experimentation is a core component of physics as a discipline. Understanding and appreciating the nature and process of experimental physics is one trademark of practicing physicists. Within the undergraduate physics curriculum, students' primary exposure to experimental physics comes from laboratory courses where students are able to directly engage with various aspects of experimentation. For many students, these courses provide their first, and potentially only, exposure to physics as an experimental science. As such, lab courses can be gateways in terms of both recruiting and retaining students within the physics major. Even 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. Building a general understanding of the process and nature of scientific experimentation is a necessary element of preparing students to be informed consumers of scientific information.

In addition to being discussed in the physics education research (PER) literature (e.g., Refs. 1 and 2), the importance of physics laboratory courses in terms of cultivating motivated and well-prepared physics graduates has also been called out by multiple professional groups.[3–5] Research focusing specifically on understanding and improving the laboratory learning environment is a new, but growing, subfield within PER. This body of work highlights the wide variety of explicit and/or implicit goals for lab courses, many of which go beyond just conveying physics concepts and extend to, for example, helping students to develop expert-like beliefs about the nature and importance of experimental physics, and science more generally.[2,3]

For physics teachers interested in improving their students' ideas about experimental physics, there is currently one research-based assessment that they can use to measure the effectiveness of their lab course at accomplishing this goal. This assessment—known as the E-CLASS (Colorado Learning Attitudes about Science Survey for Experimental Physics)—is designed to target students' beliefs about the nature and importance of experimental physics, as well as their confidence when doing physics experiments. Since its development, the E-CLASS has been administered in more than 130 physics lab courses at more than 70 institutions. The assessment has also been the subject of multiple research studies aimed at understanding and evaluating the effectiveness of various laboratory learning environments. This paper presents a summary of the E-CLASS assessment and the research that has been done using E-CLASS data with a particular emphasis on the aspects of this work that are most relevant for instructors. In doing so, we address the following questions: what is the E-CLASS? (Sec. II A), how do E-CLASS scores relate to other measures of student success in lab courses? (Sec. II B), what does E-CLASS tell us about students' ideas about experimental physics? (Sec. II C), and how do students' E-CLASS responses vary based on characteristics of the course or

Table I. List of the paper titles discussing the E-CLASS assessment which are summarized in Sec. II, along with a brief summary of the paper or key findings.

Sections	References	Title and key results/findings
II A	7	**Epistemology and expectations survey about experimental physics: Development and initial results**—Describes the motivation, theoretical grounding, development, and expert validation of E-CLASS
	10	**Students' epistemologies about experimental physics: Validating the Colorado Learning Attitudes about Science Survey for Experimental Physics**—Demonstrates the full statistical validation of E-CLASS scores for a broad student population
	11	**Alternative model for administration and analysis of research-based assessments**—Describes the development of the E-CLASS centralized and automated administration system
II B	16	**Correlating students' beliefs about experimental physics with lab course success**—E-CLASS scores show zero correlation with course grades in first-year courses, and only a small correlation with course grades in beyond-first-year courses
	17	**Impact of grading practices on students' beliefs about experimental physics**—E-CLASS scores show a moderate correlation with students' perceptions of the importance of certain activities for their grade in the course
II C	18	**Students' views about the nature of experimental physics**—Despite being relatively capable of predicting the expert-like response to E-CLASS, students often enter and leave lab courses with personal views that are inconsistent with those of experts
	19	**Improvement or selection? A longitudinal analysis of students' views about experimental physics in their lab courses**—Longitudinal increases in E-CLASS averages in consecutive courses are dominated by selection effects rather than the impact of instruction
II D	20	**Research-Based Assessment of Students' Beliefs about Experimental Physics: When is Gender a Factor?**—The gap in E-CLASS averages between men and women disappears in all but the first-year, non-physics major population after accounting for differences in preinstruction scores, course level, and major.
	29	**Impact of instructional approach on students' epistemologies about experimental physics**—Courses using research-based curricula realize higher E-CLASS scores than courses using only traditional guided labs and this effect is larger for women
	30	**Open-ended versus guided laboratory activities: Impact on students' beliefs about experimental physics**—Courses using open-ended activities realize higher E-CLASS scores than courses using only guided labs
	34	**Developing skills versus reinforcing concepts in physics labs: Insight from a survey of students' beliefs about experimental physics**—Courses focusing on developing lab skills rather than reinforcing concepts realize higher E-CLASS scores and this effect is larger for women

student? (Sec. II D). The goal of this paper is to provide a comprehensive primer on the E-CLASS assessment for laboratory physics teachers interested in learning about the assessment and its findings and/or potentially implementing it in their course. Where appropriate, we have replicated some of the quantitative results and analyses from these studies here; however, readers interested in the full quantitative analyses behind these findings summarized here should refer directly to the source publications. For reference, Table I provides a list of the papers summarized in each of the sections below along with a summary of the major findings of each.

II. THE E-CLASS ASSESSMENT

A. What is E-CLASS?

The E-CLASS assessment was developed in conjunction with laboratory course transformation efforts at the University of Colorado Boulder (CU) with the goal of providing a quantitative measure of the effectiveness of those transformations.[6,7] The choice to develop an assessment targeting students' attitudes and beliefs, rather than a more traditional conceptual assessment, was motivated by both the laboratory learning goals articulated by physics faculty and the practical constraints of laboratory courses. With respect to learning goals, researchers at CU conducted interviews and group discussions with interested physics faculty members in order to articulate consensus learning goals for CU's upper-division lab courses;[6] these goals were then used to guide the course transformation efforts and the development of an appropriate assessment instrument.[7] The resulting set of goals did not emphasize developing specific physics content or teaching specific lab techniques. Instead, faculty focused on the need for experiments that required students to apply their knowledge of physics and develop expert-like habits of mind while carrying out experiments and using measurement tools.

Another motivation for not developing a conceptual assessment for lab courses is the large variation in the content covered by these courses.[7] Research-based assessments are most useful when they can be compared across different courses. However, the specific content and experiments used in labs tend to vary significantly from course to course. Given such variation, creating a single conceptual assessment that includes content appropriate for a wide swath of even just introductory lab courses would be nearly impossible. This issue is further compounded if the goal is to design an assessment that can be used to track students' progress across multiple levels of lab courses (e.g., introductory to advanced labs). Additional motivation for the value of an attitudinal assessment like the E-CLASS comes from prior research within lecture environments, which has demonstrated that students' beliefs and expectations about the nature of doing and knowing physics can be linked to both their decision to pursue physics (i.e., retention and persistence) and their content learning in a physics course.[8,9]

An example of one of the E-CLASS's 30 Likert-style item prompts is given in Fig. 1. In each item, students are

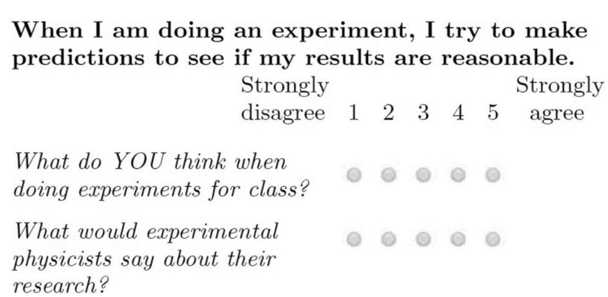

Fig. 1. An example item from the E-CLASS assessment. Students are asked to rate their agreement with the statement from their own perspective and that of an experimental physicist. See Appendix for a list of all item prompts.

presented with a statement and asked to rate their level of agreement—"strongly agree" to "strongly disagree"—from their individual point of view when doing experiments for their lab course. To ensure that they respond honestly to this first prompt, students are also asked to predict what they think an experimentalist might say about their research. This paired question format allows for comparison of students personal views with what they believe the "correct" answer to be. After its initial development, the E-CLASS was reviewed by 23 practicing experimentalists and lab instructors.[7] These expert reviews not only ensured that the assessment prompts were accurate and interesting to physics faculty but also established the consensus, expert-like response to each item (see Appendix). As described in Ref. 6, expert agreement was greater than 90% on all but six questions for which the agreement was greater than 70%. The assessment was also given to 42 students in an interview setting to ensure that students were accurately interpreting the prompts and responding in ways that were consistent with their articulated reasoning.

Scoring of the E-CLASS is done relative to the established, expert-like response for each item. While each E-CLASS item provides five possible response options from strongly agree to strongly disagree (Fig. 1), students' responses to "strongly (dis)agree" and "(dis)agree" are collapsed to a single (dis)agree category for the purposes of scoring. When reporting scores to instructors, students are awarded 1 point for each item where their response is consistent with the expert-like response and zero points otherwise. Thus, the average score on any given item represents the fraction of students who responded favorably to that question. The overall E-CLASS score is then given by a sum of the 30 individual item scores and represents the average number of favorable responses from each student. This 2-point scoring scheme does not differentiate between neutral and unfavorable responses (both receive 0 points) but provides scores with a straightforward interpretation. When scoring the E-CLASS for research purposes, neutral and unfavorable responses are not collapsed and instead are awarded 0 points and -1 point, respectively. This 3-point scoring scheme accounts for students who, for example, shift from neutral before instruction to unfavorable after instruction but provides averages with a less straightforward interpretation. Additional discussion of these different scoring schemes can be found in Ref. 10.

Students' responses to the E-CLASS are typically collected online and directly by the CU research team at the beginning and end of the lab course using a centralized and automated administration system.[11] To utilize the E-CLASS via this system, interested instructors complete a Course Information Survey (CIS, available in Ref. 12), which collects logistical (e.g., course start and end dates), structural (e.g., course content and enrollment), and pedagogical (e.g., instructional strategies) information about the course. Using information from the CIS, the system then generates unique links to the pre- and post-instruction E-CLASS for each course and sends them to the instructor to distribute to their students at the beginning and end of the laboratory component of the course, respectively. After students have responded to the survey, the system sends a list of participant names to the instructor so they may optionally offer participation credit to motivate the students to complete the assessment. After the post-instruction survey is complete, the system also generates and sends a report that breaks down the students' performance before and after instruction both overall and by item and provides comparison data from other similar courses in our national dataset. This centralized administration system reduces the burden on instructors by removing the need for them to handle most of the tasks usually associated with administering and analyzing research-based assessments.[11]

In addition to reducing the burden on lab instructors, the E-CLASS centralized system has allowed for the aggregation of a large-scale, national dataset of students' responses. To date, the E-CLASS has been administered pre- and post-instruction in more than 130 distinct lab courses at more than 70 institutions. These institutions represent a variety of different types, including 2-year and 4-year colleges, as well as masters and Ph.D. granting institutions. Courses in this dataset also span the full range of undergraduate lab course levels from introductory to advanced. This growing dataset provides a resource that can be used by instructors to compare their students' performance to students in similar courses, as well as by researchers to investigate the laboratory learning environment.

A subset of the current E-CLASS dataset was used to demonstrate the statistical validity and reliability of the E-CLASS for a broad student population.[10] This process involved demonstrating that the E-CLASS accurately and consistently measured students' beliefs about experimental physics. To do this, we calculated multiple test statistics based on students' responses from 71 courses ($N = 3591$). These data showed that, for this population of students, the E-CLASS average scores fell within acceptable ranges, adequately differentiated between high and low performing students, and provided consistent scores on subsets of items. The E-CLASS scores were also shown to be stable against retesting effects, as well as independent of testing environment (i.e., in-class vs out-of-class) and how long the students took to complete the survey. These data were also used to perform an exploratory factor analysis,[13] which showed that students' responses to the E-CLASS items did not support the existence of coherent subgroups of related questions. The results of the factor analysis imply that, despite the appeal of grouping related questions in order to facilitate interpretation of the results, students' scores on each item should be reported individually.

B. How do E-CLASS responses relate to other measures of student success?

Another important aspect of interpreting the results of a research-based assessment like the E-CLASS is determining how (or if) its scores are related to other measures of student success. However, with a few exceptions,[14,15] there is a dearth of research-based assessments specifically targeting lab courses, and those that do exist have not been widely adopted. In the absence of other research-based assessments with which to compare, another potentially interesting comparison measure of students' success is their laboratory course grade. To explore the relation between students' E-CLASS scores and their course grade, we focused on two semesters of E-CLASS data from the four core lab courses at CU spanning introductory mechanics, E&M, and modern physics, as well as upper-level electronics and optics ($N = 873$).[16] Only CU courses were included in this study because, as the authors' home institution, CU was the only institution for which we had access to matched student grade data. Final letter grades were collected for all students in these courses and were translated to standard grade point values (i.e., A = 4.0, A– = 3.7, B+ = 3.3, B = 3.0, etc.) for the purpose of quantitative analysis. We then calculated correlations between students' overall post-instruction E-CLASS score and their final courses grade point value.

We found no statistically significant correlation between students post-instruction overall E-CLASS scores and their final course grade for students in our first-year (FY) introductory lab course ($r = 0.01$).[16] Alternatively, we found a small ($r = 0.19$) but statistically significant positive correlation for students in our more advanced, beyond-first-year (BFY) courses. While we are not arguing that the relationship between E-CLASS scores and final grades is a causal one, these results do suggest that the link between course performance and students' beliefs about experimental physics is stronger in more advanced lab courses than in the FY labs. One possible explanation for this increase in correlation for BFY students was the inclusion of a final, open-ended project in some of the BFY courses. Open-ended activities that provide opportunities for student agency are one place where we might expect to see clearer connections between students' beliefs and their success or failure when completing their projects. However, even in the BFY courses, the correlation between grades and E-CLASS scores was fairly weak, suggesting that E-CLASS scores were not good predictors of students' course performance at either level or that students' beliefs were not being effectively captured by their final course grades.

To add nuance to the finding discussed above, we also examined how students' perceptions of their grades relate to their beliefs about experimental physics as measured by the E-CLASS assessment.[17] To do this, we examined students' responses ($N = 7167$) to an additional section on the post-instruction E-CLASS. This section included 23 items asking students to rate on a 5-point Likert scale how important—"not at all important" to "very important"—particular elements of experimental physics were for earning a good grade in their course. Each of these 23 questions targeting students' perceptions of grading was designed to pair with one of the E-CLASS prompts (the remaining seven E-CLASS prompts do not have a matched grade question); for example, the E-CLASS item "Working in groups is an important part of doing physics experiments" has the matched grade question "How important for earning a good grade was working in a group." Correlations between students' responses to each grade question and its matched post-instruction personal question were statistically significant for all sets of questions. Moreover, the size of this correlation was moderate ($r > 0.3$) for seven questions in both the FY and BFY student populations.[17] The magnitude of these correlations was generally larger than the correlations observed between E-CLASS scores and their actual course grades, suggesting a stronger link between students' beliefs about experimental physics and their perceptions of the grading practices than with the actual grading practices. While the causal mechanism for these correlations cannot be established from these data, these findings do suggest that laboratory course instructors should take care not only to intentionally design their grading practices to target the aspects of experimental physics they value and want their students to value but also to make those grading practices transparent to their students.

C. What does E-CLASS tell us about students' ideas about experimental physics?

The research described in Secs. II B and II A established the validity and reliability of the E-CLASS assessment, as well as how its scores relate to other measures of student success. The E-CLASS can also be used to gain insight into the status and dynamics of students' beliefs about the nature and importance of experimental physics in terms of both their personal views and their predictions of what practicing physicists would say (Fig. 1).[18] Using the full E-CLASS dataset ($N = 7167$), we calculated average scores both overall and on individual questions for students in FY and BFY courses separately (see Fig. 2). We found that undergraduate students in physics lab courses often entered and left these courses with a variety of ideas about experimental physics, some of which were inconsistent with those of practicing physicists. This trend held for both FY and BFY students, though upper-level students' personal views were more consistent with those of experts than introductory students. With respect to the impact of our lab courses in aggregate, BFY courses did not tend to result in significant shifts in students' views over the course of one semester, and introductory lab courses tended to cause small (but statistically significant) negative shifts. When predicting the responses of practicing physicists, students in upper-level courses were marginally more accurate; however, both BFY and FY students were, on average, able to correctly predict what the "expert-like" response would be for more than 80% of the questions. This finding suggests that students, even FY students with minimal experience with physics labs, are relatively good at predicting the expert response, and this trend holds even when their personal views disagree.

While the findings above suggest that BFY students tended to have more favorable views about the nature of experimental physics than FY students, it is not clear what caused this increase. For example, the increase in favorable views could be a result of instruction as students advance through the curriculum. Alternatively, it could be caused by a selection effect in which only those students whose beliefs were already expert-like persist to the higher-level lab courses. To distinguish between these two possible mechanisms, we collected E-CLASS data from three consecutive lab courses at CU over the course of eight semesters.[19] Each of these courses is required for both physics

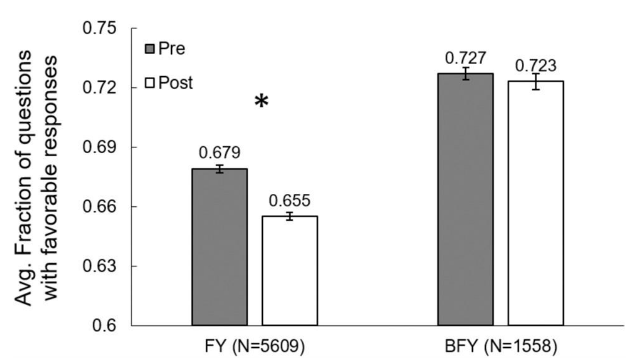

Fig. 2. Average fraction of questions with favorable responses both before and after instruction for the FY and BFY populations separately. Error bars represent one standard error of the mean and asterisks indicate statistically significant differences. Note, to facilitate visualization of the error bars, the axis has been truncated at 0.6.

and engineering physics majors at CU and was taught primarily in a traditional guided lab format. By examining the scores of students for whom we had matched pre- and post-instruction E-CLASS responses in multiple courses (i.e., longitudinal data), we found that these students both started and ended the lower level course with higher E-CLASS scores than the general population of students in that course who did not take the next course in the sequence. This finding suggests that for traditionally taught courses, the increase in E-CLASS scores between the FY and BFY courses was not driven by a cumulative impact of laboratory instruction, but rather a selection effect in which students who persisted into higher level lab courses already had higher E-CLASS scores than their peers when they started the lower level courses.

D. How do E-CLASS responses vary across courses and students?

Section II C provided a summary of general trends in students' responses to the E-CLASS both in the aggregate dataset and broken out by course level. However, we have also explored how these trends vary based on other course or student characteristics. For example, it is reasonable to expect that physics majors may have more expert-like personal views about experimental physics than students with other majors. Consistent with this expectation, we found that the population of physics and engineering physics majors had significantly higher pre- and post-instruction E-CLASS scores than the population of students with other majors (including other science majors).[10,20] Moreover, physics majors showed less negative shifts over the course of a semester than non-physics majors.[10]

Many research-based assessments in physics have also shown differences in performance between men and women.[21] To investigate the presence of gender differences in E-CLASS scores, we examined overall and by-item scores from men and women separately in the full E-CLASS dataset ($N = 7167$).[20] We found that men typically started and ended with higher E-CLASS scores than women; however, the size and even significance of this trend varied significantly depending on the subpopulation. For example, there was no significant gender gap in the population of BFY physics majors, but the gender gap did appear in the population of BFY non-physics majors. To address the intersection of these different demographic and course factors, we utilized a statistical technique known as analysis of covariance (ANCOVA).[22] ANCOVA is a statistical method for comparing the difference between population means while accounting for the variance associated with other factors. In this case, we were looking at the difference between post-instruction E-CLASS means for men and women while accounting for the impact of pre-instruction scores, major, and course level.[20] The results of the ANCOVA indicated that, after accounting for differences in pre-instruction E-CLASS score, major, and course level, gender was a significant predictor of E-CLASS performance only for FY, non-major students. This, combined with previous research linking students' beliefs and confidence with their interest and persistence in the major,[8,9] suggests that the population of FY non-physics women may be a key population for instructors and researchers to consider when working to improve students' attitudes about physics, as well as the persistence and recruitment of women into the physics major.

We also investigated the impact of differences in classroom structure or instructional strategies on students' E-CLASS responses. The majority of physics lab courses are taught using a traditional guided lab approach. However, these guided labs have been heavily critiqued as being cookbook and inauthentic.[23,24] To address this issue in introductory courses, the PER community has worked to design laboratory learning environments that allow students to engage more authentically in the process of experimental physics. Examples of these environments within our dataset include Investigative Science Learning Environment (ISLE),[25] Modeling Instruction,[26] and integrated lab/lecture environments such as studio physics[27] or SCALE-UP (Student-Centered Activities for Large Enrollment University Physics).[28]

To determine if the use of these transformed approaches to laboratory instruction had a positive impact on students' E-CLASS scores, we compared the E-CLASS responses of students in FY courses using only traditional guided labs with those of students in FY courses in which the instructor reported using one or more of these transformed instructional approaches.[29] We found that students in transformed courses had similar preinstruction scores and significantly higher post-instruction scores relative to students in purely traditional courses (Transformed: Pre = 17.1 pts, Post = 17.3 pts; Traditional: Pre = 16.4 pts, Post = 14.4 pts).[29] This post-instruction difference was driven by a negative shift in the scores of students in traditional courses and no shift in the scores of students in transformed courses. Moreover, an ANCOVA analysis of these data showed that students in transformed courses had higher E-CLASS scores even when controlling for the variance in other factors such as pre-instruction score, student major, and student gender. The ANCOVA also demonstrated that, while both men and women benefited, the increase in scores associated with transformed instructional approaches was significantly larger for women by as much as a factor of two.

Another approach to creating more authentic laboratory learning environments that better reflect actual experimental work, which has been used in both FY and BFY courses, is to incorporate open-ended activities that provide additional opportunities for students to have agency over what and how physical phenomena are investigated.[30] To investigate the impact of open-ended activities on students' E-CLASS scores, we compared the responses of students in courses using only traditional guided-lab activities to courses in which the instructor reported including one or more weeks of open-ended activities in their course. We found that courses using open-ended activities had higher pre-instruction E-CLASS scores that showed a small but statistically significant positive shift over the course of the semester (Pre = 16.9 pts, Post = 17.5 pts).[30] Alternatively, the scores of students in courses using only traditional guided-lab activities shifted down significantly over the semester (Pre = 16.2 pts, Post = 14.4 pts).[30] To account for the difference in pre-instruction score, we also used an ANCOVA, which showed a positive impact from including open-ended activities even after controlling for the variance associated with course level, student gender, and student major.

As previously discussed, there are a wide range of implicit and explicit goals for lab courses.[2,3,31] In addition to fostering students' understanding of, and appreciation for, the nature and importance of experimental physics,[2,6,31,32] two common goals are to reinforce the physics concepts taught within the lecture courses[31–33] and to develop students' practical lab skills.[3,6] These learning goals may not be independent, meaning that a focus on one goal may help or hinder the achievement of one of the other two. To explore this dynamic, we asked instructors to report whether the main purpose of their lab course was to reinforce physics concepts (concepts-focused), develop lab skills (skills-focused), or both (both-focused).[34] BFY courses were much more likely than FY courses to report being skills-focused, while almost no BFY courses reported being concepts-focused. For FY courses, we found that students in skills-focused courses showed more expert-like post-instruction responses and more favorable shifts than students in either concepts-focused or both-focused courses and these differences were statistically significant (Skills: Pre = 16.9 pts, Post = 17.6 pts; Concepts: Pre = 17.7 pts, Post = 14.9 pts).[34] We also found that instructors in skills-focused courses used fewer verification labs, provided more opportunities for student agency, and more often asked students to quantify uncertainty in a measurement and maintain a lab notebook than instructors in concept-focused courses. An ANCOVA also demonstrated that, while both men's and women's scores were higher in skills-focused courses, this effect was significantly larger for women.

III. SUMMARY AND RECOMMENDATIONS

Undergraduate physics lab courses represent an important and unique element of the undergraduate physics curriculum. In particular, lab courses can be one of the primary environments in which students gain experience with the nature and process of experimental physics, as well as its place within the discipline as a whole. The E-CLASS assessment was designed as an easy-to-use tool to assess physics lab courses with respect to how successful they are at encouraging students to adopt beliefs about the nature and importance of experimental physics that are consistent with those of practicing physicists. This paper presents a comprehensive overview of this laboratory-focused, research-based assessment that can be used to guide and assess attempts to improve our physics lab courses. The E-CLASS was developed based on consensus learning goals for lab courses articulated by physics faculty. Additionally, the statistical validity and reliability of E-CLASS scores was demonstrated using a large dataset of students' responses from multiple courses and institutions.

Using this large dataset, we also established that, while post-instruction E-CLASS scores have at most a small correlation with students' final course grades, these scores correlate more strongly with students' perceptions of the grading structure, specifically what elements of experimental physics they believe are important for earning a good grade in the course. This dataset has also been extensively studied to determine overall trends in students' ideas about experimental physics, as well as how these trends vary across different student populations, courses, and types of instruction. These data also represent a resource for physics lab instructors and researchers looking to compare their students' performance against other courses with similar structure or content.

For the physics lab instructor interested in improving their students' beliefs about the nature and importance of experimental physics, the results summarized above suggest a number of potential starting places and things to consider. For first-year, introductory courses, use of one of the research-based instructional approaches that have been developed for these courses (i.e., ISLE, Modeling instruction, or integrated lab lecture) resulted in higher average E-CLASS scores and more favorable shifts. These transformed instructional approaches have the added benefit of being well documented and established in the literature, reducing the need for instructors to come up with completely new materials or instructional approaches. Additionally, first-year courses that focused more on developing students' practical lab skills than on reinforcing physics content also saw higher average E-CLASS scores and more favorable shifts.

In both of the cases discussed above—using transformed instructional approaches and focusing on developing lab skills—the bump in these first-year students' E-CLASS scores was larger for women than for men. This trend is particularly important given that analysis of gender gaps in E-CLASS scores indicated that first-year women, particularly non-physics majors, represent the most at risk population with respect to the appearance of significant gender gaps. Thus, the use of transformed instructional approaches and/or the focus on skills development have the potential not only to improve students E-CLASS scores but also to reduce the gender gap amongst first-year students.

Another strategy that our data suggest that lab instructors might use in both first-year and beyond-first-year courses is the inclusion of some open-ended activities. Open-ended activities are those activities that provide opportunities for students to take some agency in what and how physical phenomena are investigated. For all levels of lab courses in our data, the inclusion of one or more weeks of open-ended activities resulted in higher average E-CLASS scores and more favorable shifts. One advantage of this approach is that it can potentially be implemented with less drastic changes to the lab environment. Many traditional guided-lab activities can be modified to be more open-ended through relatively small scale changes to the course framing and lab guides without requiring a complete overhaul of the actual lab experiments.

There are several important caveats and considerations for a physics lab instructor to take into account when considering the recommendations above. The E-CLASS assessment targets only a particular subset of the various potential goals for a physics lab course; there are many other valuable learning outcomes that E-CLASS does not target. Instructors should carefully consider whether the items on the E-CLASS align with their specific learning goals prior to using the assessment. Additionally, our large dataset allows us to detect overall trends in students' scores based on different instructional strategies; however, this large dataset also tends to wash out the variation in the overall success of any particular implementation of these strategies. The particulars of how each strategy is implemented will likely have a significant impact on its overall success. None of the recommendations given above represents a magic bullet for improving a lab course; successful implementations will need to take into account the goals and constraints of the particular courses, as well as the specific needs of the student population.

ACKNOWLEDGMENTS

This work was funded by the NSF-IUSE Grant No. DUE-1432204 and NSF Grant No. PHY-1125844. Particular thanks to the members of PER@C for all their help and feedback.

APPENDIX: LIST OF E-CLASS ITEM PROMPTS

Consensus expert-like response given in parentheses at the end of each statement.

Q1: When doing an experiment, I try to understand how the experimental setup works. (agree)

Q2: If I wanted to, I think I could be good at research. (agree)

Q3: When doing a physics experiment, I don't think much about sources of systematic error. (disagree)

Q4: If I am communicating results from an experiment, my main goal is to create a report with the correct sections and formatting. (disagree)

Q5: Calculating uncertainties usually helps me understand my results better. (agree)

Q6: Scientific journal articles are helpful for answering my own questions and designing experiments. (agree)

Q7: I don't enjoy physics experiments. (disagree)

Q8: When doing an experiment, I try to understand the relevant equations. (agree)

Q9: When I approach a new piece of lab equipment, I feel confident I can learn how to use it well enough for my purposes. (agree)

Q10: Whenever I use a new measurement tool, I try to understand its performance limitations. (agree)

Q11: Computers are helpful for plotting and analyzing data. (agree)

Q12: I don't need to understand how the measurement tools and sensors work in order to carry out an experiment. (disagree)

Q13: If I try hard enough, I can succeed at doing physics experiments. (agree)

Q14: When doing an experiment, I usually think up my own questions to investigate. (agree)

Q15: Designing and building things is an important part of doing physics experiments. (agree)

Q16: The primary purpose of doing physics experiments is to confirm previously known results. (disagree)

Q17: When I encounter difficulties in the lab, my first step is to ask an expert, like the instructor. (disagree)

Q18: Communicating scientific results to peers is a valuable part of doing physics experiments. (agree)

Q19: Working in a group is an important part of doing physics experiments. (agree)

Q20: I enjoy building things and working with my hands. (agree)

Q21: I am usually able to complete an experiment without understanding the equations and physics ideas that describe the system I am investigating. (disagree)

Q22: If I am communicating the results from an experiment, my main goal is to make conclusions based on my data using scientific reasoning. (agree)

Q23: When doing an experiment, I try to make predictions to see if my results are reasonable. (agree)

Q24: Nearly all students are capable of doing physics experiments if they work at it. (agree)

Q25: A common approach for fixing a problem with an experiment is to randomly change things until the problem goes away. (disagree)

Q26: It is helpful to understand the assumptions that go into making predictions. (agree)

Q27: When doing an experiment, I just follow the instructions without thinking about their purpose. (disagree)

Q28: I do not expect doing an experiment to help my understanding of physics. (disagree)

Q29: If I don't have clear directions for analyzing data, I am not sure how to choose an appropriate analysis method. (disagree)

Q30: Physics experiments contribute to the growth of scientific knowledge. (agree)

[1]Valerie K. Otero and David E. Meltzer, "100 years of attempts to transform physics education," Phys. Teach. **54**(9), 523–527 (2016).

[2]Ricardo Trumper, "The physics laboratory—A historical overview and future perspectives," Sci. Educ. **12**(7), 645–670 (2003).

[3]AAPT Committee on Laboratories, "AAPT recommendations for the undergraduate physics laboratory curriculum", 2015.

[4]Joint Task Force on Undergraduate Physics Programs, "Phys21: Preparing physics students for 21st-century careers", 2016.

[5]Steve Olson and Donna Gerardi Riordan, "Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics," Report to the president, 2012.

[6]Benjamin M. Zwickl, Noah Finkelstein, and H. J. Lewandowski, "The process of transforming an advanced lab course: Goals, curriculum, and assessments," Am. J. Phys. **81**(1), 63–70 (2013).

[7]Benjamin M. Zwickl, Takako Hirokawa, Noah Finkelstein, and H. J. Lewandowski, "Epistemology and expectations survey about experimental physics: Development and initial results," Phys. Rev. Spec. Top. PER **10**(1), 010120-1–14 (2014).

[8]David Hammer, "Epistemological beliefs in introductory physics," Cognit. Instr. **12**(2), 151–183 (1994).

[9]Laura Lising and Andrew Elby, "The impact of epistemology on learning: A case study from introductory physics," Am. J. Phys. **73**(4), 372–382 (2005).

[10]Bethany R. Wilcox and H. J. Lewandowski, "Students' epistemologies about experimental physics: Validating the colorado learning attitudes about science survey for experimental physics," Phys. Rev. Phys. Educ. Res. **12**, 010123-1–11 (2016).

[11]Bethany R. Wilcox, Benjamin M. Zwickl, Robert D. Hobbs, John M. Aiken, Nathan M. Welch, and H. J. Lewandowski, "Alternative model for administration and analysis of research-based assessments," Phys. Rev. Phys. Educ. Res. **12**, 010139-1–8 (2016).

[12]"E-CLASS website," 2015, <tinyurl.com/ECLASS-physics>.

[13]Norm O'Rourke, R. Psych, and Larry Hatcher, *A Step-by-Step Approach to Using SAS for Factor Analysis and Structural Equation Modeling* (Sas Institute, Cary, NC, 2013).

[14]James Day and Doug Bonn, "Development of the concise data processing assessment," Phys. Rev. Spec. Top. Phys. Educ. Res. **7**, 010114-1–14 (2011).

[15]Seshini Pillay, Andy Buffler, Fred Lubben, and Saalih Allie, "Effectiveness of a gum-compliant course for teaching measurement in the introductory physics laboratory," Eur. J. Phys. **29**(3), 647–659 (2008).

[16]Bethany Wilcox and H. J. Lewandowski, "Correlating students' beliefs about experimental physics with lab course success," in *Physics Education Research (PER) Conference*, College Park, MD, 2015, pp. 367–370.

[17]Bethany Wilcox and H. J. Lewandowski, "Impact of grading practices on students' beliefs about experimental physics," in *Physics Education Research (PER) Conference 2017*, Cincinnati, OH, 2017.

[18]Bethany R. Wilcox and H. J. Lewandowski, "Students' views about the nature of experimental physics," Phys. Rev. Phys. Educ. Res. **13**, 020110 (2017).

[19]Bethany R. Wilcox and H. J. Lewandowski, "Improvement or selection? a longitudinal analysis of students' views about experimental physics in their lab courses," Phys. Rev. Phys. Educ. Res. **13**, 023101 (2017).

[20]Bethany R. Wilcox and H. J. Lewandowski, "Research-based assessment of students' beliefs about experimental physics: When is gender a factor?," Phys. Rev. Phys. Educ. Res. **12**, 020130-1–11 (2016).

[21]Adrian Madsen, Sarah B. McKagan, and Eleanor C. Sayre, "Gender gap on concept inventories in physics: What is consistent, what is inconsistent, and what factors influence the gap?," Phys. Rev. Spec. Top. Phys. Educ. Res. **9**, 020121-1–15 (2013).

[22]Albert R. Wildt and Olli Ahtola, *Analysis of Covariance* (Sage, Newbury Park, CA, 1978), Vol. 12.

[23]National Research Council, *Adapting to a Changing World–Challenges and Opportunities in Undergraduate Physics Education* (The National Academies Press, Washington, DC, 2013).

[24]Carl Wieman, "Comparative cognitive task analyses of experimental science and instructional laboratory courses," Phys. Teach. **53**(6) 349–351 (2015).
[25]Eugenia Etkina and Alan Van Heuvelen, "Investigative science learning environment: Using the processes of science and cognitive strategies to learn physics," in *Physics Education Research Conference (PER) 2001*, Rochester, NY, 2001.
[26]Malcolm Wells, David Hestenes, and Gregg Swackhamer, "A modeling method for high school physics instruction," Am. J. Phys. **63**(7), 606–619 (1995).
[27]Jack M. Wilson, "The cuple physics studio," Phys. Teach. **32**(9), 518–523 (1994).
[28]Robert J. Beichner, Jeffery M. Saul, Rhett J. Allain, Duane L. Deardorff, and David S. Abbott, "Introduction to scale-up: Student-centered activities for large enrollment university physics," in *Proceedings of the 2000 Annual Meeting of the American Society for Engineering Education, Session 2380*, Seattle, WA, 2000, pp. 1–13.
[29]Bethany Wilcox and H. J. Lewandowski, "Impact of instructional approach on students' epistemologies about experimental physics," in *Physics Education Research (PER) Conference 2016*, Sacramento, CA, 2016, pp. 388–391.
[30]Bethany R. Wilcox and H. J. Lewandowski, "Open-ended versus guided laboratory activities: Impact on students' beliefs about experimental physics," Phys. Rev. Phys. Educ. Res. **12**, 020132-1–8 (2016).
[31]Susan R. Singer *et al.*, *America's Lab Report: Investigations in High School Science* (National Academies Press, Washington, DC, 2006).
[32]Robin Millar, *The Role of Practical Work in the Teaching and Learning of Science* (National Academy of Sciences, Washington, DC, 2004).
[33]Carl Wieman and N. G. Holmes, "Measuring the impact of an instructional laboratory on the learning of introductory physics," Am. J. Phys. **83**(11), 972–978 (2015).
[34]Bethany R. Wilcox and H. J. Lewandowski, "Developing skills versus reinforcing concepts in physics labs: Insight from a survey of students' beliefs about experimental physics," Phys. Rev. Phys. Educ. Res. **13**, 010108-1–9 (2017).