

PAPER • OPEN ACCESS

Assessing students' views about experimental physics in a German laboratory course

To cite this article: E. Teichmann *et al* 2024 *J. Phys.: Conf. Ser.* **2750** 012013

View the [article online](#) for updates and enhancements.

You may also like

- [Spectrally resolved wave-mixing between near- and far-infrared pulses in gas](#)
M Clerici, D Faccio, L Caspani et al.
- [Signatures of dark excitons in exciton-polariton optics of transition metal dichalcogenides](#)
Beatriz Ferreira, Roberto Rosati, Jamie M Fitzgerald et al.
- [Close Encounters with Heisenberg: uncertainty in secondary school](#)
L Galante, M Arlego, M Fanaro et al.



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

Assessing students' views about experimental physics in a German laboratory course

E. Teichmann¹, H. J. Lewandowski^{2,3} and M. Alemani¹

¹ Physics and Astronomy Institute, University of Potsdam, Potsdam 14476, Germany

² Department of Physics, University of Colorado, Boulder, Colorado 80309, USA

³ JILA, National Institute of Standards and Technology and University of Colorado, Boulder, Colorado 80309, USA

E-mail: alemani@uni-potsdam.de

Abstract. Physics laboratory courses (PLC) have been recently the topic of several research studies examining their effectiveness at reaching their goals. As a result, a discussion about the effectiveness of traditional PLC for students' content knowledge, skills, and "expert-thinking" acquisition has developed. Critical for the investigation of students learning in those settings has been the development of research-based assessments tools. An example of those is the Colorado Learning Attitudes about Science Survey for Experimental Physics (E-CLASS). Recently, we translated the E-CLASS into German and set up a centralized survey administration system for instructors, allowing data acquisition and automated data analysis. Previously, we described this process and presented the preliminary results of the study of the introductory PLC at the University of Potsdam (UP). Here, we present an extended study that allows us to make stronger conclusions about students' views about experimental physics at the UP. Overall, we find that students at US institutions have a higher level of "expert-like" views than students at the UP.

1. Introduction

Physics laboratory courses (PLC) are an essential part of the physics curriculum. They offer students a unique opportunity to engage with the processes of experimental physics and acquire several important skills for their future careers. They also have the potential to enhance students' identity as physicists at an early stage and align students' views with those of expert experimental physicists.

Recent research studies, however, have demonstrated the shortcomings of so called "concept-focused" and "cook-book" laboratory courses, which aim to improve conceptual knowledge through prescriptive lab activities, for the acquisition of experimental skills and for development of students' expert-like views on the nature of experimental physics [1, 2, 3].

Several efforts have been done in recent years to improve students learning in PLC and several PLC around the world have been reconstructed using research-based approaches [4, 5, 6, 7, 8]. Crucial for these efforts was the development of research-based assessment tools [9, 10, 11]. Those tools allow instructors to easily assess their courses along different dimensions and critically and carefully guide course transformations based on the assessment results. An example of these research-based assessment tools is the Colorado Learning Attitudes about Science Survey



for Experimental Physics (E-CLASS). It is widely used in physics laboratory classes to assess students' views and attitudes about experimental physics [10, 3].

In a previous publication, we presented how we translated the E-CLASS into German and how we validated the translation [12]. We also described how we set-up a centralized system for instructors for data acquisition and automated analysis. This system was developed in accordance with the European data privacy standards. We called this German translation of E-CLASS the GE-CLASS, where the 'G' stays for German. In our former publication, we also presented the first example use of the GE-CLASS to study the introductory laboratory courses at the University of Potsdam (UP). The results of this study are very important as they represent the first study of a German laboratory course using the GE-CLASS [12]. The number of students represented in the dataset was, however, limited. In order to overcome this limitation, we continued to collect data with the GE-CLASS at UP. In this publication, we present the analysis of this larger data set. This allows us to make stronger conclusions about the assessment of the introductory PLC at UP using the GE-CLASS and about the comparison between the E-CLASS results from US institutions and the results obtained at the UP.

In the following, we will first briefly introduce the E-CLASS survey, then present the details about the data sets and describe the methods for the data analysis. After that, we will present and discuss the results of the data analysis including a comparison between the GE-CLASS results and the one obtained at US institutions.

2. The E-CLASS

The E-CLASS is composed of 30 items each related to a different aspects of experimental physics[10]. The variety of dimensions of experimental physics addressed by the 30 items makes this survey widely applicable even for courses with different learning goals. Instructors can, in fact, focus on different subsets of items that are directly related to their learning goals. Each item of the survey has a core statement regarding a different aspect of experimental physics. A list of all statements can be found in the appendix. For each core statement, students first rate on a five-point Likert scale (from strongly disagree to strongly agree) their personal agreement to the core statement. Secondly, they rate what they think a practicing experimental physicist would answer using the same Likert scale. We refer to the first kind of questions as "YOU-questions" and to the second kind as "EXPERT-questions". Students' answers for both the YOU- and EXPERT- questions are then compared to the answers given by a group of 23 practicing experimental physicists. We call the answers given by those group of experts as the expert reference (ER). By comparing students' answers before and after the laboratory course with the answers given by the ER, it is possible to evaluate the impact of instruction upon students' views.

3. Data collection and Analysis Methods

The GE-CLASS data set presented here was collected from 2019 to the end of 2022. In this time we obtained 300 valid responses.

To match student responses from pre- to post-instruction for the GE-CLASS, each student is given a course ID code and then they create a self-generated anonymous code. We described the creation of this code in details in reference [12]. A survey response is considered valid only if the course ID code and the self-generated anonymous code match for the pre- and post- survey. We also used a control question (asking students to answer with a particular given answer) to make sure students were reading the questions before answering. An incorrect answer for the control question would cause that survey response to be invalid, and, thus, not used in the analysis.

Among the 300 valid responses obtained, 191 responses are from courses for the first year (FY) and 109 are from courses beyond the first year (BFY) (see Table I). The introductory laboratory

instruction for physics major students at the UP takes place during the first four semesters of the curriculum, and for each of those semesters, the laboratory course is mandatory.

In 2016, we started to transform our introductory courses from “concept-focused” to “skill-focused”. During this transformation, we defined a set of skills important for experimentation (e.g., modeling, design, and communication) and created activities designed for students to practice a particular set of skills. As a result of such transformation, each semester’s laboratory course has specific learning goals (built upon each other) and settings. We also had set as broader goal for our course transformation to align students’ personal views of experimental physics to that of experts. The details about the course types, their goals and settings are described in detail in reference [12]. The data used for this study were collected from 26 different instances of those PLC conducted between the first and the fourth semesters at the UP.

Table 1. Details about the data used in this study. The GE-CLASS data were collected entirely at the UP. All students were physics major students. The E-CLASS data are a subset of the larger data set. We use data from only physics majors. We randomly sample a portion of the E-CLASS data such that we have the same ratio of first year (FY) and beyond first year (BFY) students as in the GE-CLASS data.

	GE-CLASS	E-CLASS
Institution	1	90
FY courses	14	306
BFY courses	12	282
Matched students responses FY	191	1073
Matched students responses BFY	109	611
Total students responses	300	1684

For the comparison between the GE-CLASS and E-CLASS results, we have selected a subset of data from the entire E-CLASS data set. Note here that the entire data set has been recently made publicly available in an anonymous form [13]. From the larger E-CLASS data set, we considered only physics major students and we randomly sample a portion of the E-CLASS data such that we have the same ratio of FY and BFY students as in the GE-CLASS data. Note that about 95% of the E-CLASS data set was collected at U.S. institutions. Therefore results from the E-CLASS data set can be considered representative of what is happening in US institutions. The details of the resulting data sets for the E-CLASS and GE-CLASS are shown in Table I.

For the data analysis, we compared students’ responses to the survey to the responses given by the ER. For this, we first reduced the five-point Likert scale to a three-point Likert scale, by collapsing “strongly (dis)agree” and “(dis)agree” into a single category. We then assigned a numerical score of +1 for agreement with the ER, 0 for neutral, and -1 for disagreement with the ER.

4. Results

Using the above described analysis method, we looked at students’ responses to the YOU-questions item-by-item before and after instruction. In figure 1, we show the agreement with the answers of the ER for the E-CLASS data (above) and for the GE-CLASS data (below). The results for the pre/post instruction are represented for the E-CLASS using gray squares

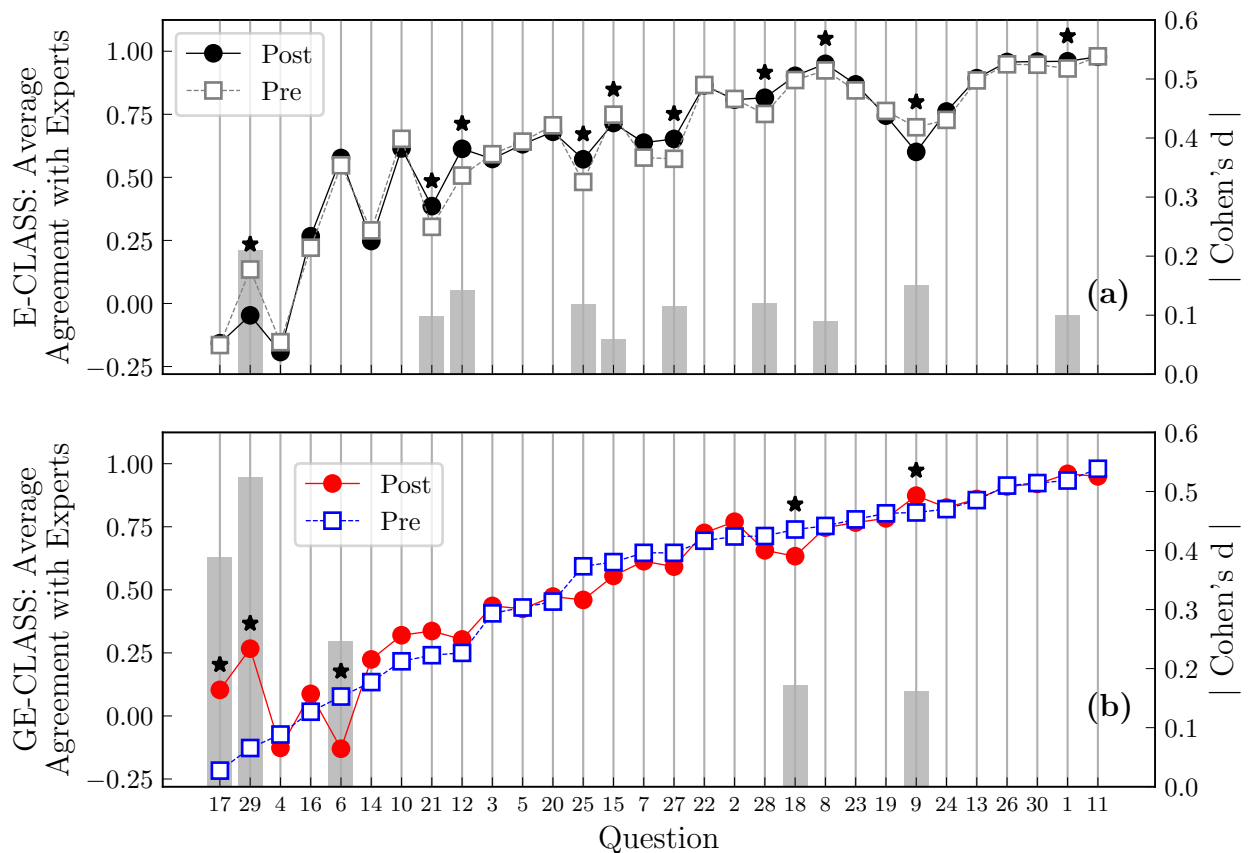


Figure 1. Item-by-item agreement with the ER for the YOU questions. E-CLASS data are in graph (a), while the GE-CLASS data are in graph (b). Stars indicate pre/post changes that are statistically significant as calculated with the Mann-Whitney U-test[14]. The absolute values of the Cohen's d are indicated in the figure as bars and can be read on the right axis.

and black circles respectively and for the GE-CLASS using blue squares and red circles. The items in the figure have been ordered based on the level of students' agreement with the ER in the GE-CLASS pre-instruction data. In the figure, we indicate statistically significant changes upon instruction with the black stars. For testing the statistical significance, we used the non-parametric Mann-Whitney U-test [14] with 95% confidence level and with the null hypothesis that the two samples are coming from the same population. To evaluate the practical significance of statistically significant shifts, we calculate an effect size using Cohen's d [16]. The Cohen's d values are shown in the figure 1 as grey bars and their values can be read on the right axis.

Looking at the GE-CLASS data in figure 1(b), we can see that instruction has a positive impact on items 17 (*If I don't have clear directions for analyzing data, I am not sure how to choose an appropriate analysis method*), 29 (*When I encounter difficulties in the lab, my first step is to ask an expert, like the instructor*) and 9 (*When I approach a new piece of lab equipment, I feel confident I can learn how to use it well enough for my purposes*) with medium to small effect sizes. On the other hand, we observe statistically significant negative shifts for items 6 (*Scientific journal articles are helpful for answering my own questions and designing experiments*) and 18 (*Communicating scientific results to peers is a valuable part of doing physics experiments*).

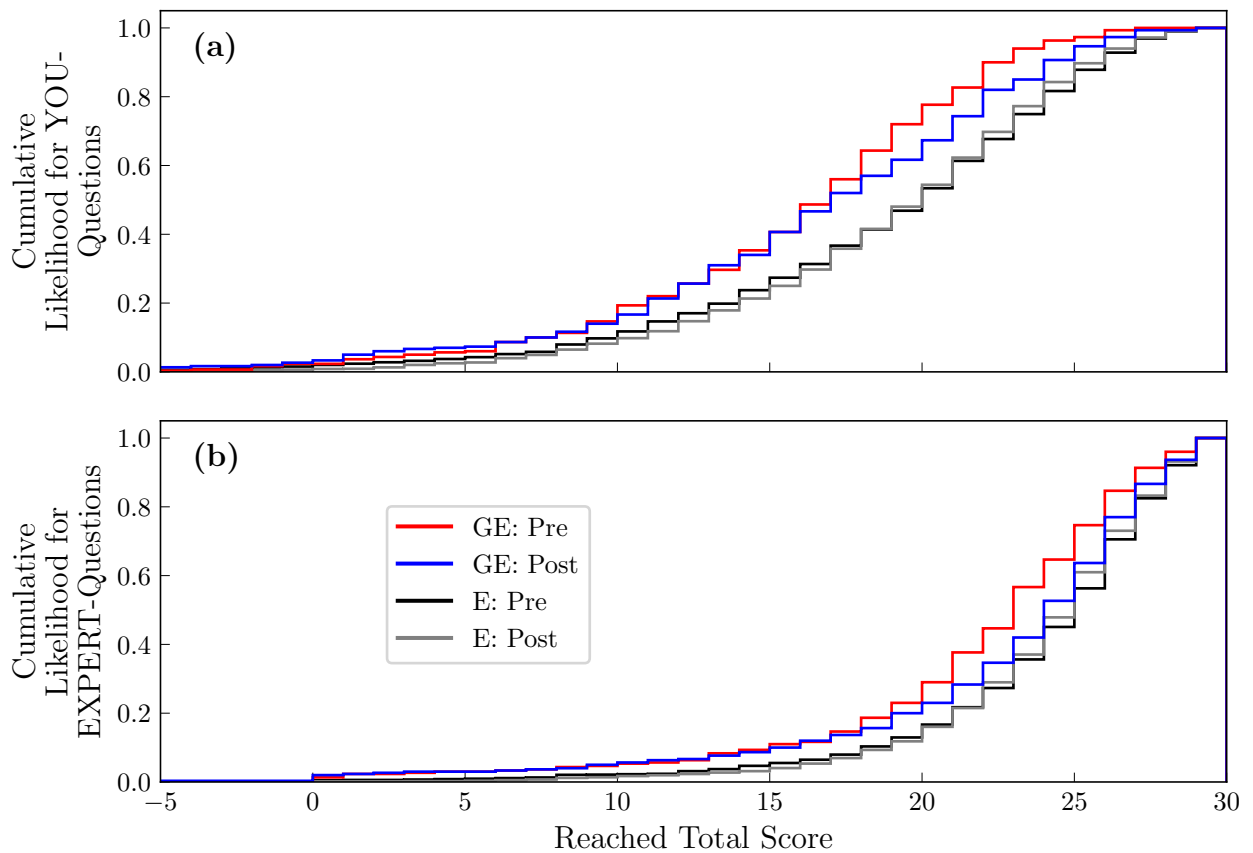


Figure 2. Cumulative likelihood distributions of the agreement with experts as a function of the total score reached in the survey (obtained by summing up on all items). In (a) are the results for the YOU-questions, in (b) are the results for the EXPERT-questions. Red and blue lines represents GE-CLASS results for the pre- and post- surveys respectively. Black and gray lines correspond to the pre- and post E-CLASS results.

There are several statistically significant changes upon instruction for the E-CLASS data, some positive and some negative, but the effect sizes of those changes are always small. Overall, we notice a higher level of agreement with the ER for the E-CLASS data than the GE-CLASS.

We see that the four items with the lowest and largest agreement with the ER are the same for the GE-/E-CLASS. Both groups score highest and lowest (i.e., have “expert-like” and “non-expert-like” thinking) on items regarding the same aspects of experimental physics independently on how different the two educational environments are. Items in between the extremes have different order of agreement with the ER for the two data sets.

To further analyse our data, we plotted in figure 2 the cumulative likelihood distribution of the reached total scores. This corresponds to the integral of the fraction of students that reached up to a certain total score in the survey. Note here that the maximum total score a student can reach is +30 (when a student responds to all statements as the ER), while the minimum is -30 (when a student responds opposite to that of the ER for all statements). The best possible distribution would be zero everywhere with a sharp peak at +30. This would, in fact, mean that all students have responded to all statements as the ER. The cumulative likelihood distribution

of the scores for the YOU-questions are shown in figure 2(a) and for the EXPERT-questions in figure 2(b). GE-CLASS data are in red and blue (for pre- and post- surveys respectively), while E-CLASS data are in black and grey respectively.

The E-CLASS distributions are shifted to the right with respect to the GE-CLASS (see figures 2(a) and (b)). To investigate if the observed differences between the E-CLASS and GE-CLASS cumulative distributions are statistically significant, we used the non-parametric Anderson-Darling k-samples statistical test [15]. For this test, we used as null hypothesis that the two- samples are drawn from the same population and considered a confidence level of 95%.

We analysed all four cases for the comparison including the difference between:

- (i) E- and GE-CLASS distributions of the YOU-questions in the pre-survey
- (ii) E-/GE-CLASS distributions of the YOU-questions in the post-survey
- (iii) E-/GE-CLASS distributions of the EXPERT-questions in the pre-survey
- (iv) E-/GE-CLASS distributions for the EXPERT-questions in the post-survey.

We found that the two samples of the E- and GE-CLASS do not originate from the same distribution ($p\text{-value} \ll 0.05$) in all those four cases. This means that the E-CLASS distributions are more “expert-like” than the GE-CLASS case.

Moreover, as we found in our previous study, in both GE- and E-CLASS cases, the distributions for the EXPERT-questions in figure 2(b)) are shifted to the right, i.e. are more “expert-like”, with respect to the distributions for the YOU-questions (in figure 2(a)).

When considering changes upon instruction for the GE-CLASS distribution only, we observe statistically significant changes between GE-CLASS pre- and post-distributions for both YOU- and EXPERT-questions. On the other hand, the shifts upon instruction of the E-CLASS data (between pre- and post-) are not statistically significant for both YOU- and EXPERT-questions.

Finally, we found for YOU-questions in the GE-CLASS case (see figure 2(a)) that positive effects upon instruction happens mostly for students with already high levels of expert-views before instruction, while students with low levels of expert-views worsen their views. The opposite tendency is true for the E-CLASS results.

5. Discussion

The results presented indicate that students at US institutions have a higher level of “expert-like” views with respect to students at the UP. Notice here that preliminary results using the GE-CLASS at other German institutions, which will be published elsewhere, suggest us that this result is not specific to the UP. If this preliminary result is confirmed after collecting a larger data set, we will be able to conclude that more effort needs to happen at German universities and high schools for aligning students views about experimental physics to those of experts, which is also true in the US [17].

We also found that in both educational environments (i.e., UP and US) students have an internal contradiction between their personal views and their view of experts, which has been explored in more depth in the US context [17]. They know what experts think, but do not practice these beliefs while they do experiments.

Moreover, the results presented in figure 2 show that the effects upon instruction for GE-CLASS data are larger than for the E-CLASS data. We observe changes towards more “expert-like” views but, as discussed in section 4, we need to attend more to students that start with less “expert-like” views, as, unfortunately, these students’ views tend to become more novice-like upon instruction.

We can examine the impact of instruction in further detail by looking at which items show a positive or negative shift. We found in the case of the GE-CLASS that students score better after instruction for items related to students confidence in solving problems and making decisions (items 17, 29 and 9). This result is rewarding, as we put a lot of effort into restructuring our

course to encourage students to make decisions independently while trying to support them with our course scaffolding. The negative shift observed for question 6 (*Scientific journal articles are helpful for answering my own questions and designing experiments*) is unsurprising, as this aspect of experimental physics is not a learning goal in our laboratory course and therefore there are no related activities included in the course. This finding shows that students align their views about experimental physics in accordance with the activities in the course. Importantly, GE-CLASS and E-CLASS both ask students about their views with respect to their lab class and not experimental physics in general. Thus, students may come into the course expecting these components to be a part of experimental physics, but then do not engage in the activity in their course, which could lead to a negative shift. The negative shift of item 18 (*Communicating scientific results to peers is a valuable part of doing physics experiments*) was also found in our previous study [12]. Since it is a learning goal of our course, we have started to put effort into creating activities that demonstrate the importance of peer communication while performing experiments. Until now, however, we have experimented in only a couple of cases with new ways to include this aspects in the PLC, but have not managed to include those ways on a regular basis in our PLC.

6. Conclusions

In conclusion, we have assessed the introductory PLC at the UP, using the German version of the E-CLASS and compared the results at the UP with those obtained at US institutions. The insights gained from this study are important feedback in the iterative process of improving laboratory instruction. We will work to improve how students engage in scientific communication in particular. This is a central goal for the UP course. Overall, the lower level of “expert-like” views at the UP compared to US institutions indicates the need for a stronger focus on epistemology.

Acknowledgments

We would like to thank M. Path and R. Zumpe for their assistance during data acquisition during the laboratory courses. We also acknowledge support from the Stifterverband and the Baden-Württemberg Stiftung (Project No. H1205228500830925) and by National Science Foundation (Grant No. PHY-1734006).

Appendix A. List of the E-CLASS questions

We show in the parenthesis at the end of the E-CLASS items the answer of the ER. 'A' stands for AGREE and 'D' stands for DISAGREE.

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

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

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

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

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

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

Q7: I don't enjoy doing physics experiments. (ER: D)

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

- 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. (ER: A)
- Q10: Whenever I use a new measurement tool, I try to understand its performance limitations. (ER: A)
- Q11: Computers are helpful for plotting and analyzing data. (ER: A)
- Q12: I don't need to understand how the measurement tools and sensors work in order to carry out an experiment. (ER: D)
- Q13: If I try hard enough I can succeed at doing physics experiments. (ER: A)
- Q14: When doing an experiment I usually think up my own questions to investigate. (ER: A)
- Q15: Designing and building things is an important part of doing physics experiments. (ER: A)
- Q16: The primary purpose of doing a physics experiment is to confirm previously known results. (ER: D)
- Q17: When I encounter difficulties in the lab, my first step is to ask an expert, like the instructor. (ER: D)
- Q18: Communicating scientific results to peers is a valuable part of doing physics experiments. (ER: A)
- Q19: Working in a group is an important part of doing physics experiments. (ER: A)
- Q20: I enjoy building things and working with my hands. (ER: A)
- Q21: I am usually able to complete an experiment without understanding the equations and physics ideas that describe the system I am investigating. (ER: D)
- Q22: If I am communicating results from an experiment, my main goal is to make conclusions based on my data using scientific reasoning. (ER: A)
- Q23: When I am doing an experiment, I try to make predictions to see if my results are reasonable. (ER: A)
- Q24: Nearly all students are capable of doing a physics experiment if they work at it. (ER: A)
- Q25: A common approach for fixing a problem with an experiment is to randomly change things until the problem goes away. (ER: D)
- Q26: It is helpful to understand the assumptions that go into making predictions. (ER: A)
- Q27: When doing an experiment, I just follow the instructions without thinking about their purpose. (ER: D)
- Q28: I do not expect doing an experiment to help my understanding of physics. (ER: D)
- Q29: If I don't have clear directions for analyzing data, I am not sure how to choose an appropriate analysis method. (ER: D)
- Q30: Physics experiments contribute to the growth of scientific knowledge. (ER: A)

References

- [1] Wieman C and Holmes N G 2015 *Am. J. Phys.* **83** 972
- [2] Volkwyn T S, Allie S, Buffler A, and Lubben F 2008 *Am. J. Phys.* **4** 010108
- [3] Wilcox B R and Lewandowski H J 2016 *Phys. Rev. ST Phys. Educ. Res.* **12** 010123
- [4] Etkina E, Heuvelen A V, White-Brahmia S, Brookes D T, Gentile M, Murthy S, Rosengrant D, and Warren A 2006 *Phys. Rev. ST Phys. Educ. Res.* **2** 020103
- [5] Karelina A and Etkina E 2007 *Phys. Rev. ST Phys. Educ. Res.* **3** 020106
- [6] Zwickl B M, Hu N F D and Lewandowski H J 2015 *Phys. Rev. ST Phys. Educ. Res.* **11** 020113
- [7] Dounas-Frazer D R and H. J. Lewandowski 2018 *Eur. J. Phys.* **39** 064005

- [8] Holmes N G and Wieman C 2018 *Phys. Today* **71** 38
- [9] Day J and Bonn D 2011 *Phys. Rev. ST Phys. Educ. Res.* **7** 010114
- [10] Zwickl B M, Hirokawa T, Finkelstein N, and Lewandowski H J 2014 *Phys. Rev. ST Phys. Educ. Res.* **10** 010120
- [11] Walsh C, Quinn K N, Wieman C, and Holmes N G 2019 *Phys. Rev. Phys. Educ. Res.* **15** 010135
- [12] Teichmann E, Lewandowski H J and Alemani M 2022 *Phys. Rev. Phys. Educ. Res.* **18** 010135
- [13] Aiken J M and Lewandowski H J 2021 *Phys. Rev. Phys. Educ. Res.* **17** 020144
- [14] Mann H B and Whitney D R 1947 *Ann. Math. Stat.* **18** 50
- [15] Scholz F W and Stephens M A 1987 *J. Am. Stat. Assoc.* **82** 918
- [16] Cohen J 1988 *Statistical Power Analysis for the Behavioral Sciences* (Lawrence Erlbaum Associates, Hillsdale, NJ)
- [17] Wilcox B R and Lewandowski H J 2017 *Phys. Rev. Phys. Educ. Res.* **13** 020110.