Development of a global landscape of undergraduate physics laboratory courses

Gayle Geschwind,[1,2] Micol Alemani,[3] Michael F. J. Fox,[4] P. S. W. M. Logman,[5,6] Eugenio Tufino,[7] 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, 390 UCB, Boulder, Colorado 80309, USA*
[3]*University of Potsdam, Institute of Physics and Astronomy, Karl-Liebknecht-Straße 24-25, 14476 Potsdam, Germany*
[4]*Department of Physics, Imperial College London, Prince Consort Road, London, SW7 2AZ, United Kingdom*
[5]*Leiden Institute of Physics, Leiden University, Leiden, The Netherlands*
[6]*Leiden University Graduate School of Teaching, Leiden University, Leiden, The Netherlands*
[7]*Department of Physics, University of Trento, Trento, Italy*

(Received 24 April 2024; accepted 19 August 2024; published 20 September 2024)

Physics education research (PER) is a global endeavor, with a wealth of work performed at a variety of institutions worldwide. However, results from research into undergraduate physics laboratory courses are often difficult to compare due to the broad variations in courses. We report here how we developed and validated a survey to classify these courses, as well as compare and contrast them. This will be useful in two key endeavors: comparisons between PER studies and providing useful data for individual instructors hoping to improve their courses. While we are still in the process of collecting sufficient data to create a full taxonomy of laboratory courses, we present here details of the survey creation itself, including its face, construct, and content validation, as well as a first look at the data collected, which includes a broad landscape of lab courses in 41 countries. We used both quantitative and qualitative methods to analyze the data collected. Some of these results include similarities between courses, such as students often using preconstructed apparatuses and instructors hoping for students to learn technical skills. We also find differences in courses, such as in the number and types of goals of the course, as well as the activities students participate in. Thus, this survey and its results can provide information relevant to both researchers and instructors.

DOI: 10.1103/PhysRevPhysEducRes.20.020117

I. INTRODUCTION

Physics education is a global endeavor. As we strive to find the best methods of education for the next generation of physics students, undergraduate physics education can benefit from an international perspective, both for improving and comparing courses and to aid students studying worldwide. In today's world, international collaboration is growing due to more accessible technology and the need to engage scientists across the world to answer important questions and solve critical issues that are global in nature [1]. Thus, physics education should cross the boundaries of countries to form a cohesive structure to enhance the education of future scientists. In order to best conduct physics education research to improve education, we first need to understand the similarities and differences in how physics is taught worldwide, including degree requirements, classroom environments, and experiences of students. This will also help future collaborations among physicists, as they can better understand their collaborators' previous educational experiences, which could lead to a better appreciation of the variety of backgrounds of participants in a collaboration.

Here, we focus on the context of undergraduate physics laboratory courses due to the significant current work in this space [2–5], as well as the importance of physics laboratory courses in general and the unique skills they can provide for students [6].

Our ultimate goal is to create a taxonomy, or classification scheme, of undergraduate physics laboratory courses that can be applied worldwide. This taxonomy would have numerous applications, including gathering information about courses so that lab instructors and course developers may be inspired by others, as well as facilitating comparisons that may be made through physics education research (PER) studies. From a research perspective, it is difficult to know whether studies done in certain courses can be compared to others as lab courses are a rich and complex space with a large variety of implementations. A taxonomy could help classify these courses, so research results can be used appropriately without overgeneralization. Thus, an important step in the process of improving physics laboratory education is to understand what these courses currently look like globally.

A taxonomy could also be used to standardize comparison data that are currently presented in reports to instructors about their courses from research-based assessment instruments (RBAIs) [7,8], assessments used to determine how well students collectively meet course learning goals (as opposed to assessments used to individually evaluate students) [9]. Typically, results from RBAIs for laboratory courses present instructors with both their own students' performance and data from other courses for comparison. Unfortunately, these comparison data usually include data from all students who have taken the assessment, regardless of the type of course it was used in. This makes it difficult to know whether the comparison data are appropriate. A taxonomy could be used to select comparison data from only similarly characterized classes.

On the practical side, a taxonomy could allow instructors to learn from one another about what they do in order to improve or transform their courses. One could also use a taxonomy of lab courses to learn about student experiences in different systems; this may be helpful in graduate admissions and in facilitating international collaborations among physicists with different educational backgrounds.

Due to the scale required to achieve the creation of a taxonomy of lab courses, we decided to create a survey to learn about such courses. Hence, the specific goals of this paper are as follows:

1. Provide initial steps related to the development and validation of a survey designed to capture information about undergraduate physics lab courses worldwide.
2. Report initial results from the survey to demonstrate the diversity of lab courses that the survey is able to capture.

Here, we present a detailed view of the development of the survey, as well as information to support claims of validity. Specifically, we discuss construct validity, content validity, and face validity of the survey. Construct validity aims to determine if the survey measures the concept it is intended to; in our application, we analyze whether the survey addresses information about undergraduate physics lab courses. Content validity aims to determine if the survey is fully representative of what it aims to measure, which in this case means that we would like to determine if the survey fully spans the space of relevant information about lab courses. Finally, face validity determines whether the content of the survey is suitable for its goals—that is, if the questions in the survey appear to measure information regarding these courses [10–17]. Further, one particular aspect of face validity relevant due to the international nature of this survey is to ensure the survey is interpretable for both native and non-native English speakers.

We also present a first look at the similarities and differences across all survey items from 217 lab courses in 41 countries represented in the initial data collections. In doing so, this also supports the first goal of this work, by providing evidence to support content validity in demonstrating the items capture a wide range of responses.

Future efforts will work to collect significantly more survey responses to be able to apply clustering methods [18] to achieve the ultimate goal of creating a taxonomy of lab courses that can be used by both instructors and education researchers.

II. BACKGROUND

A. Prior PER on laboratory courses with a global perspective

Although undergraduate laboratory courses offer unique opportunities for students to learn experimental skills, such as experimental design and data analysis techniques, they have historically been overlooked as less important than their lecture counterparts in the curriculum [19] and have become only recently a focus of research [3].

Currently, the field of PER in laboratory courses is focused largely in the United States but is quickly becoming more international. To illustrate the diversity of research on lab courses and, therefore, the need to develop a taxonomy that will allow us to compare both studies and courses, we highlight some contributions to the PER literature from the community outside of the United States. For example, one paper from Taiwan discusses methods of integrating technology into physics lab courses, including potential options for virtual and remote laboratory instruction, including the creation of a framework for others hoping to use technology to support inquiry-based activities [20]. A recent paper from Finland discusses the modernization of a physics lab course at the University of Helsinki, with details about shifting to more open-inquiry activities [21]. Other work from Finland details methods of assessing students' work in lab courses, including different types of examination and feedback [22]. Recent work from India discusses shifting an undergraduate electronics lab toward open-ended activities in order to help improve students' research skills, including technical skills, problem-solving abilities, and collaboration [23]. One paper from Germany presents a similar shift from prescriptive laboratory activities toward a skills-based course with more authentic experiments, finding that students are more engaged and are better able to master important laboratory skills, such as keeping a lab notebook [24]. Further, in Italy, a group of researchers examined the transformation of a lab course to include activities with Arduinos and smartphones, including an open-ended aspect [25]. Other work, from the Netherlands, shows a tendency toward open-inquiry lab courses [26–28].

In addition to these efforts in in-person labs, the COVID-19 pandemic prompted research into remote laboratory activities. One such paper from the Netherlands described this abrupt transition where they investigated the use of Arduinos with open-inquiry activities [29]. In France, Arduinos have been used in a project-based lab course for third-year students, in which they are given complete decision-making control over their experimental setup and what to investigate with the Arduinos in order to help them learn more about the nature of experimental physics [30]. Further work from the Netherlands reports the creation of a Mach-Zehnder interferometer from children's toys with an Arduino detector. This work highlights the ability to achieve experimental physics learning goals without the need for expensive resources [31].

In addition to research from individual countries, collaborations between researchers in different countries have become more common as modern technology facilitates connections with people around the world. For example, a collaboration between universities in Finland and the United States investigated students' abilities in critical thinking over the course of a semester [32]. Another collaboration involving researchers in Germany, Finland, Switzerland, and Croatia examined the use of digital experiments in physics lab courses, including the development of a questionnaire in four languages to investigate student use of these online experiments aimed at remote learning [33]. Researchers in Germany and the United States investigated student views of experimental physics in German lab courses, finding distinct differences between these students and their American counterparts. [34]. Separately, researchers in Denmark, Czechia, and Slovenia explored teacher education regarding lab courses with global input from a discussion at a conference. This work specifically focused on learning goals and the role of labs in teaching physics [35].

These collaborations and international studies can reveal similarities and differences in the ways that lab courses are taught around the world [36], allowing us to challenge our perspectives on how lab courses are taught. However, it is difficult to directly compare research from all of these studies without an understanding of the basic laboratory course structure, goals, and activities at these different institutions. Even within the United States, laboratory courses differ vastly between different colleges and universities [37]; adding an international component to the mix that further complicates this. Some prior research has investigated physics lab instruction in North America [37] using course instructor surveys from RBAIs, but this does not include a broader international aspect. Holmes *et al.* found that physics lab instruction across North America varies considerably in many aspects, including course goals, activities, and pedagogical methods [37]. Another recent paper compared instructional strategies during the pandemic from one university from each of the three countries involved (United States, Sweden, and Australia) and found that, despite the widely varying locations and cultural constructs, all universities struggled with successfully implementing emergency remote instruction [38].

One class of studies that aims to be generalizable is the use of RBAIs. These assessments are intended for gathering information about students' performance collectively, rather than for assigning individual students' grades. They are frequently used to determine whether a course is meeting its learning goals [9]. One such RBAI is the Colorado Learning Attitudes about Science Survey for Experimental Physics (E-CLASS) [39–42]. Despite the name, this assessment instrument has been used broadly with greater than 100,000 total responses from many countries (mostly in North America, Europe, and Asia). It has also been translated into several different languages (including Swedish [43], Italian [25], German [34], Chinese, Spanish, Hebrew [44], Norwegian, and Amharic) to allow for easier administration in other countries. Another globally used RBAI, the Physics Lab Inventory of Critical Thinking (PLIC) [45–48], which focuses on critical thinking in physics labs, including questions related to data analysis and measurement uncertainty. This survey is available in Chinese, Finnish [32], German, and Spanish, in addition to English. Both surveys collect information about the type of lab course the RBAI is being administered in, however, they are both U.S.-centric, which partly motivates the need for the present work.

B. Prior work on characterizing undergraduate physics courses

There has been some previous work on characterizing undergraduate physics courses that we can build on for the current study.

Prior research has worked to characterize physics theory courses, especially at the introductory level. For example, some research focuses on introductory physics problems and creation of a taxonomy of the cognitive processes required to solve these problems [49,50]. This taxonomy can be applied to help develop assessments in courses, guide curriculum development, and determine learning goals for introductory courses. However, little research has been done to apply such a taxonomy to higher-level theory courses or to compare theory courses between different institutions.

Some recent research has focused on lab courses. One effort to characterize laboratory courses is related to the administration of E-CLASS and PLIC. To implement E-CLASS using the centrally administered version in English, instructors complete a course information survey, which gathers information about the course itself, including information about the level (introductory or beyond introductory), whether it is algebra or calculus based, the number of students and staff, and how frequently students participate in various activities. A nearly identical set of questions is used in the administration of the English version of PLIC [37]. While this information is useful in many research applications, it is limited in scope and does not provide enough for course classification for worldwide applications [37,51]. For example, it is missing information about activities, pedagogies, and course design. Some questions from the course information surveys for E-CLASS and the PLIC were used as input for the survey discussed in this paper, including questions about the level of the course (introductory versus beyond introductory), type of institution, and information about course goals [37], though these current instruments lack important details about courses, hence the need to develop a new instrument.

Another effort we draw from was initiated during the COVID-19 pandemic, where research was conducted regarding the switch to emergency remote laboratory instruction, which helped us establish content and face validity. Researchers created a survey to collect information from course instructors about the structure, goals, and components, as well as other features of their remote or hybrid lab courses [52–54]. The pandemic instructor survey included some similar questions to those in the E-CLASS course information survey but also included questions about changes to course goals and activities (whether they were incorporated in the course both before the pandemic and while the course was remote). While the data were helpful in analyzing differences between instruction pre-pandemic versus during the pandemic, the questions were not extensive enough to fully characterize laboratory courses to create a global taxonomy of such courses and were created locally on an extremely short timescale to capture the rapidly evolving situation.

III. METHODS

This section discusses the various methods used in the paper. These methods relate to the creation of the survey, the validation of the survey, and data collection. First, we present information about the survey's creation, including the initial brainstorming steps and organization into a logical survey format, followed by interviews with instructors contributing to the establishment of construct, content, and face validity of the survey. Finally, we discuss the dissemination of the survey broadly, including the limitations of our data collection.

A. Survey creation and validation

1. Initial creation of ideas and organization

The project idea emerged from a small workshop (about 20 attendees) at Imperial College London in April 2022, where PER researchers who had used the E-CLASS survey in Europe and Western Asia were invited to present. From discussions among attendees, it became apparent there was not a clearly defined way to compare lab courses from different countries. A subset of attendees decided to embark upon a project to create a lab taxonomy that would be broadly applicable. The idea would be to create and validate a survey to collect data on lab courses around the world and use that data to create a taxonomy. The description of this initial creation of the survey will provide evidence for both face and content validity.

The first step toward the development of the survey was a collective brainstorming session over Zoom, where we discussed the most important aspects of undergraduate physics laboratory courses and created a virtual whiteboard to collect and partially organize the ideas. Many of the ideas organized on the whiteboard, which eventually morphed into the various categories in the survey, were based on the Spinnenweb (spiderweb) representation of curriculum and learning proposed by van den Akker [55]. This model describes which different facets play a role in students' learning. A large fraction of the survey was modeled off this structure, including sections such as grouping, assessment, goals, activities, content, instructional staff roles, and materials and resources. This brainstorming session and collection of ideas was the first step of content validation, as the authors attempted to ensure that the survey would cover all aspects of lab courses. The authors' experiences working in experimental physics and lab-focused PER (see Appendix A) contribute to the content of the survey being fully representative of what we aim to measure, with no important concepts missing.

After this initial process, the results of the whiteboard activity were loosely organized into a document by category. This document included many of the ideas that eventually went into the survey but was missing several important concepts and included many items that we decided to remove. For example, a list of equipment that students might have access to in the course was removed for being too unwieldy for both survey participants and researchers to handle, as well as being outside the scope of the survey, thus ensuring that content validity was maintained. We also added questions from previous E-CLASS/PLIC and pandemic instructor surveys. Face validity was determined by the authors during the initial preparation of the survey questions, as the questions all related to the goal of determining information about lab courses such that a taxonomy could be created.

The list of lab course components that emerged was then transformed into a survey format through an iterative process with feedback provided by all authors, based on personal experience and the literature, and a small number of physics education researchers separate from the project team. This involved organizing the information we wished to probe into meaningful categories, ordering those categories in an intuitive way, and turning ideas we wanted to probe into questions. Many of the iterations of the survey involved language changes, with interviews (as detailed below) later validating the wording choices made for the

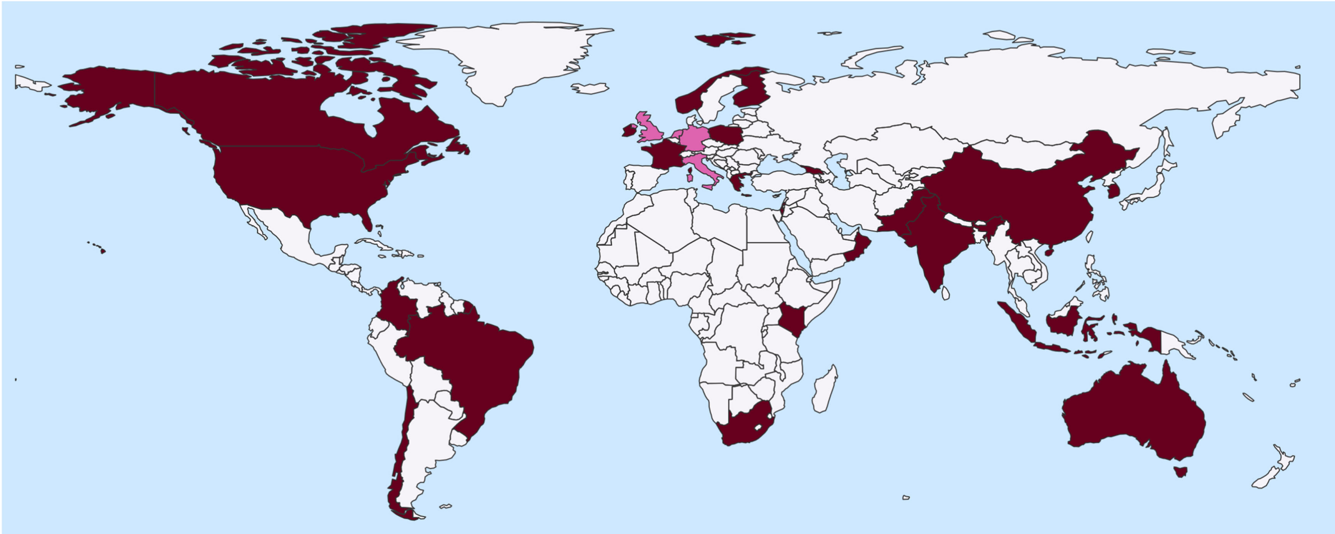

FIG. 1. World map indicating the countries where we interviewed lab instructors. Each maroon country had one instructor interviewed, aside from the United States, which had two. The pink countries represent the home countries of the authors of this paper, aside from the United States, and therefore no interviews were solicited from instructors in these countries because the authors could provide the necessary feedback. We interviewed 23 instructors from 22 countries. A full list of these countries can be found in the Appendix.

survey. At this stage, the survey document was coded into a Qualtrics survey in order to prepare for the interview validation phase of survey development.

2. Interviews with lab instructors

In order to further validate the survey, we performed interviews with lab instructors. These interviews provide evidence for three types of validity: construct validity, content validity, and face validity, as described in the introduction of this paper. In this section, we will describe the process of the interviews and how the questions asked relate to establishing the different aspects of validity discussed.

Interviews were solicited from contacts known by the authors. These contacts were compiled into a list and solicitations were chosen such that only one person per country would be contacted and there would be a reasonable worldwide spread of participants. We chose to exclude Germany, Italy, the Netherlands, and the United Kingdom from interviews due to already having researchers on our team from these countries. We did include the United States in interviews due to the diversity in universities in this country [56–60]. We conducted interviews only with those who currently teach or have previously taught undergraduate physics laboratory courses.

In total, we conducted 23 interviews with participants from 22 countries; the United States was sampled twice. A map of the countries participating in interviews is shown in Fig. 1 and a list of these countries can be found in the Appendix. We contacted 32 lab instructors from 26 countries in total, leading to approximately a 72% response rate. Interviews were solicited in several rounds, and we determined after 23 interviews that changes to the survey had become minimal and therefore, further interviews would be unnecessary, showing content and construct validity. We made changes to the survey after nearly every interview and presented the newest version to the next interviewee.

Author G. G. conducted all of the interviews over Zoom. Each interview lasted between 33 and 76 min. Because the entire survey is in English, the interviews were also conducted in English. While this limits the pool of both interviewees and those who can participate in the survey, and certainly places limits on some countries more than others, translating the survey is currently outside of the scope of this project. Both video and audio of the interviews were recorded for later analysis. Additionally, the interviewer took notes during the interview and, in most cases, these notes provided the basis for further changes to the survey.

Interviewees were provided with a link to the survey during the interview and were asked to talk through the questions while sharing their screens. They were instructed to consider all undergraduate physics laboratory courses they had knowledge of when considering the survey questions. We probed whether (i) the survey captures the construct of what a lab course is in order to determine construct validity, (ii) that the items and response options cover all aspects of lab courses in order to determine content validity, and (iii) that the wording of the questions is unambiguous in order to determine face validity.

An interview protocol was created to validate each of the survey questions by answering three validation questions:

1. Are the survey questions understandable to those who are not from a country represented by an

author? Are the survey questions interpreted in the way they were intended?
2. Do the survey questions make sense in the context of courses the instructor has taught?
3. Do the survey questions fully span the space of information that the instructor feels is important to capture about undergraduate physics laboratories?

For most survey questions, interviewees were asked whether they understood the question, as well as whether anything was missing; in some cases, interviewees were also asked to explain in their own words the concepts and ideas presented in certain questions, especially in the goals and activities sections of the survey.

In the first validation question, we probe whether the (mostly) American English used throughout the survey is understandable both to those who have learned English as a second language and to those who speak English as their primary language but use a different dialect (British, Canadian, or Australian English, for example). Differences arise due to variations in terms used by British English (e.g.,“revise” means both “study” and “alter” in British English but only means alter in American English), as well as terms not commonly used for those who speak another language as their primary language (e.g., the term “rubric” was determined to be unfamiliar to many non-native English speakers, but when the concept was described, interviewees were familiar with it). This first validation question captures construct validity to ensure that instructors taking the survey understand the questions as intended and therefore, the survey measures what it is supposed to. Further, this helps ensure face validity as we can be certain that the questions are suitable for the goals of the survey.

To answer the second validation question, interviewees were asked whether each survey question was applicable in courses familiar to them, thus providing evidence of construct and face validity. Here, we gather evidence of construct validity because we can ensure the questions measure what we are intending them to (i.e., information about their courses), rather than having questions on the survey that may not be relevant. Further, face validity is present here because we are also able to determine that the questions are all related to the ultimate goal of the survey; interviewees noted that none of the questions seemed to be outside the scope of the survey.

In some cases, questions had to be added or logic had to be introduced in order to ensure that people taking the survey would be able to appropriately answer all questions without confusion due to certain questions not being applicable to their courses. For example, in one case, we added a question about whether the lab course meets weekly. In the case that it does, the instructor is sent to a page with the original questions about number of hours per week the course meets and the number of weeks the course runs. In the case that the course does not meet weekly, instructors are sent to a different set of questions in which they are asked to describe their course meetings (how often, how many hours per term, etc.).

Answering the third validation question helped us to add, remove, and revise questions to ensure we gathered all of the information we need to create a taxonomy and describe the state of undergraduate physics laboratory courses around the world, thus providing evidence of content validity. Interviewees were asked whether questions regarding the goals of the course and activities students participate in were complete or missing important relevant information. Many times, interviewees suggested additions that helped to make the survey more accurately capture the full breadth of components of the courses, thus providing evidence for face validity such that the questions are all related to the goal of the survey, as well as ensuring content validity as this ensures that both the survey and each individual question fully spanned the space of information we hoped to gather about that specific topic being probed. Interviewees were also asked at the end of the survey whether they felt anything was missing that they thought should have been included in order to ensure content validity.

Many changes to the survey occurred concurrently with the interview process. As interviewees made comments about the survey, changes were discussed with members of the team and implemented so that updated questions could be validated with subsequent interviewees. Details of the results of this process are discussed in Sec. IVA as we go through each item on the survey. By the end of the interview process, we had evidence of construct, content, and face validity because the interviewees were no longer suggesting significant changes (additions, subtractions, etc.) to the questions or to the entire survey itself. In terms of construct validity, interviewees were able to describe the meaning of the questions with definitions as intended by the authors and answer these questions in the context of the courses they teach. For content validity, interviewees had no more additions to suggest to the survey in order to fully capture their courses by the end of this process. Finally, in terms of face validity, interviewees expressed that the questions were all relevant to their courses and the authors believed these questions were all related to the goals of the survey.

B. Survey dissemination, data collection, and data analysis

We first disseminated the final version of the survey using all of the international contacts of the authors. We compiled a list of our contacts (∼130) and then emailed a solicitation to take the survey to people we knew through our professional networks. This solicitation invited people to take the survey, as well as to pass it along to others they know (whether in their own department or at other institutions), which is often referred to as snowball sampling [61,62]. We included snowball sampling as it is typically used for difficult-to-access populations. In this case, physics lab instructors are such a population as their positions and titles vary and contact information is often not readily available, as previously discussed in another context [63].

We also posted the survey on a variety of listservs relevant to lab instructors. These included the ALPhA listserv [64], two American Physical Society (APS) discussion boards (Forum on Education [65] and Topical Group on Physics Education Research [66]), a newsletter distributed to members of GIREP, and a JISCmail forum for physics education researchers and instructors in the United Kingdom [67]. Further, the survey was advertised during the ALPhA Beyond First Year (BFY) IV Conference, American Association of Physics Teachers Summer Conference, Physics Education Research Conference, and the GIREP Conference (all in July 2023) during various authors' talks and poster presentations.

Next, we compiled a list of 171 countries worldwide and focused our efforts on those not represented in our sample thus far, especially those within world regions not well represented. Each author then searched for publicly available information about institutions with physics departments (e.g., from institution websites) within 10–15 unique countries and contributed to a database with contact information for department heads and lab instructors. This approach has limitations, as many institutions did not have contact information easily accessible. However, this led to a large increase in responses: we received about 53 additional responses from 24 countries to add to our dataset. The solicitation email we sent to these institutions requested that they send to their contacts as well. In selecting countries, we chose places where we did not have a well-represented sample, such as much of Asia, Eastern Europe, South America, and Africa. This helped spread the survey to a more global audience than just where our contacts were located.

Data presented here were collected from 20 June 2023 until 10 January 2024, though the survey is still open and collecting responses [68]. We include data only from people who responded to a minimum number of questions (about 80% of the survey). This removed 121 responses with no questions answered and 18 responses that answered only some questions, leaving 217 out of the initial 356 responses. Because we do not force responses to most questions on the survey, aside from a couple that are necessary for future logic within the survey, each question has a varying number of responses. The number of respondents is reported for each part of the results section as appropriate. The median time to take the survey was 18 min.[1] We present descriptive statistics of the survey responses in order to answer our second research question, as well as to provide evidence that our survey is able to capture a wide variety of courses from multiple countries.

Finally, in analyzing the data, we compared responses to two questions, one regarding the goals of the course and another regarding the activities students participated in during the course in order to determine if these activities align with the course goals. Similarly, we also matched these course goals with items that might be used to evaluate students for their final course grades. Authors G. G. and H. J. L. along with an outside PER postdoctoral researcher worked together to match the course goals with activities, as well as the course goals with items graded. We then determined whether instructors are connecting their course goals with the activities students participate in and in the ways they are evaluated in the course. As a simple example, the course goal "Developing lab notebook keeping skills" can be matched with activities "maintain an individual lab notebook" and "maintain a group lab notebook." This goal can be matched with "Lab notebooks" under items graded. After G. G., H. J. L., and the outside researcher came to an agreement about matching, an analysis was done to determine how many activities instructors chose that matched the course goals, as well as how many items were graded.

1. Limitations of the data collection methods

First, aside from the United States, the countries where the authors are located (Germany, Italy, the United Kingdom, and the Netherlands) are oversampled based on the number of institutions present in each country, while most other countries are undersampled by this same metric. For example, we have only three responses from China and two responses from India, two largely populated countries with strong physics programs. We also have only four responses from Africa and nine responses from South America, therefore undersampling large areas of the world. We are also missing many countries entirely. The United States is underrepresented compared to responses received from other countries based on the number of higher education institutions—we have only 63 responses for the United States. Additionally, within the United States, there are a large variety of types of institutions [56–60], so getting a truly representative sample of institutions would be difficult and is something we do not currently have at this point in our data collection.

Second, some of our responses are clustered at specific institutions. One example of this is Canada—while we have six responses from Canada, five of them are from one university within Canada and all six responses are from a single province.

Third, since the survey was disseminated primarily using our contacts and listservs we are members of, those who responded are more likely to be interested or involved in

[1]Median is reported to exclude the outliers of those who leave the survey open for multiple days without actively filling it out before submitting it, thereby skewing the mean and making it an inappropriate statistic to report.

physics education research to some extent. This also biased the countries from which we received responses, as those with relationships with the authors were more likely to fill out the survey and pass it on to their colleagues, hence why the country bias is skewed toward our home countries. There is also bias in who chose to fill out the survey: those with an interest in improving their programs and are invested in the quality of lab teaching are more likely to fill out the survey.

Fourth, the survey is available only in English. While many of our colleagues do speak English, we miss many people who do not know English well enough to complete the survey. We chose not to translate the survey at this time due to constraints on resources and expertise.

Finally, because of these limitations, we do not present uncertainties in our tables in most cases. Due to the sampling bias in our data, our unknown systematic error is likely larger than the uncertainty determined by statistical means. It could be therefore misleading to present the statistical uncertainty.

Considering all of the limitations of our dataset, we advise caution when interpreting results. While the development of the survey and validating it are not subject to these limitations, the data addressing our second goal of this paper—providing an overview of undergraduate physics lab courses globally—are affected by them. In particular, strong trends are likely to be applicable, but weaker trends and subtle differences shown in the data may not accurately represent the full global landscape of labs.

IV. RESULTS AND DISCUSSION

A. Lab taxonomy survey and validation

This study was designed to create and validate the final version of the survey; the validation, as well as the results from the survey itself, is presented here as major results of this work. The survey was developed through a systematic process and extensive interviews, and will hopefully serve as a foundational tool for years to come. An adaptation of the Qualtrics version of the survey is presented in Supplementary Material [69].

As described previously, we looked for evidence of the validity of the survey along three axes: face, construct, and content validity. Face validity was determined to be upheld when no further adjustments to the survey were required during the interview phase.

Next, construct validity was also determined during survey development. The ultimate goal of the project is the creation of a taxonomy of courses; however, we do not have enough data to create this scheme yet and, therefore, do not have all of the evidence required for full construct validity. Nevertheless, the interviews did provide evidence that this type of validity is present in our survey, as the questions measured information about their courses according to the interviewees.

Finally, we determine content validity primarily through the interviews. Here, the goal of determining whether any questions were missing from the survey was our main method of providing content validity. During the interview phase, we determined the survey has content validity using the criteria described in Sec. III.

The survey is delivered online via Qualtrics and remains open for data collection [68]. We note that the survey does not collect information about the instructors (e.g., name, email address, and demographics) but rather focuses only on the course itself.

The survey is structured with eight separate sections to capture a variety of information including overall course and institution characteristics, students, grouping of students, instructional staff, goals, activities, evaluation, and an optional open text box for any additional items, including a request for lab activity titles.

Interviews helped shape the final form of the survey. The first general category of survey edits were simple improvements, including the addition of a progress bar, bolding "select all that apply" wherever it appeared in order to draw attention to it (based on several interviewees missing this text in the questions), inclusion of a back button, and other general readability improvements. These were minimal edits that did not significantly change the contents of the survey but rather improved the user experience. In addition to these changes, minor wording changes and clarifications were made throughout the survey to improve understandability to a wider audience. We also note that for each section, the number of questions did not change significantly during interviews but rather the content and wording of the questions did.

Separately, we found that instructors might benefit from taking this survey, as well as from published works that will come from it. Some of the questions, especially those about branded approaches to instruction and the use of RBAIs, were very interesting to interviewees. These questions include links to outside resources about various instructional methods and assessments. One interviewee said:

> It looks very interesting, actually… I'm going to open all of them… that looks nice. Here, we are very far away from that… I'm going to come back to this. Thank you so much for that [sic] links,

while examining the question about branded approaches to instruction (such as ISLE and SCALE-UP). This interviewee was excited to learn about methods they had not previously been familiar with. Questions such as these, and those about course goals and activities, can allow instructors to reflect on their courses and consider possible new teaching methods, assessments, and course goals. We hope this unintended impact of the survey can also be useful for improving laboratory instruction.

In the following, we present each section of the survey together with a description of the changes we made during their development.

1. Overall characteristics

This section of the survey asks for basic information both about the institution where the course is taught and about the course. Institution questions include the location and name of the institution and highest degree it grants. Course information collected in this section includes general information, such as the name of the course, the intended level of the course (introductory vs beyond introductory), a checklist of physics topics covered by the course, the number of students enrolled in a typical term, and the basic setup of the course—for example, whether students participate in project work, and the types of experiments students do (weekly, many experiments per week, or experiments that last longer than one week). Other questions include whether the lab course is integrated with a lecture course and whether the course includes lectures on statistics, data analysis, or experimental techniques. This section also includes questions about project work which are displayed if participants indicate that projects are part of the course.

Many edits were made to this section during the interview phase. Several items were added to the list of topics that might be covered in a physics laboratory course, including quantum information, geophysics, and modern physics; this change occurred due to interviewees suggesting topics that they felt were important but were not included in the original list. A question was also added to this category to probe the level of the lab, introductory or beyond introductory, after it became clear that the year of the students taking the course is not sufficient to provide this information (e.g., a course for life science majors might be mostly second- or third-year students, but it might be the first physics laboratory course these students take and is therefore considered to be introductory). Another question was added to probe whether the course meets weekly. If the respondents choose “no,” we added an open text box for them to provide details of their course meeting schedule. We added this text box as several interviewees mentioned that their courses followed unique course meeting schedules. Because we can not account for every such case, we determined that an open text box was the best method to collect this information and may help refine future iterations of the survey.

2. Students

This section of the survey first asks about students’ majors and the percentage of students in the course earning a degree in physics or astrophysics. Finally, this section asks participants to estimate how many years the students have been at the university.

The most significant change that occurred in this section due to interviews was to include a question asking about the percentage of students in the course that are physics majors. We decided that simply asking about the majors of students was not enough information to determine whether the course is intended for physics majors, nonmajors, or both; the inclusion of this question helps make that more clear. In addition to this change, more majors were added to the list of potential degrees students might be earning as a result of interviewee requests.

3. Students’ group work

This section of the survey inquires about how students work in the lab: alone or with others. If students work with others, participants are asked several follow-up questions, including the typical size of a group, how groups are chosen, and whether students stay in the same groups for the entire course.

No significant changes were made to this section during the interview phase aside from slight wording edits.

4. Instructional staff

This part of the survey asks about the types of instructional staff present in the lab with students. These might include faculty, lab technicians, graduate teaching assistants (TAs), and undergraduate learning assistants (LAs). In this section, participants are also asked about training provided to TAs and LAs—both the frequency of this training and the topics covered (e.g., familiarization with equipment, pedagogy instruction, and grading training).

In this section, we added lab technicians during the interview phase at the request of interviewees. Slight wording changes were also made to the questions inquiring about the types of training.

5. Goals

In this section, many potential goals for a physics laboratory course are listed, and participants rank these on a Likert scale consisting of Major Goal, Minor Goal, Not a Goal, and Future Goal (not currently a goal). The development of this unique Likert scale is discussed in more detail below.

The list of goals is as follows:

- Reinforcing physics concepts previously seen in lecture (confirming known results or seeing theory in an experiment).
- Learning or discovering physics concepts not previously seen in lecture.
- Developing technical knowledge and skills (e.g., making measurements and hands-on manipulation of equipment).
- Designing experiments.
- Developing mathematical model(s) of experimental results.

- Learning how to analyze and interpret data (e.g., linear regressions and uncertainty).
- Learning how to visualize data (e.g., plotting).
- Developing lab notebook keeping skills.
- Developing scientific writing skills (e.g., lab reports).
- Developing other communication skills (e.g., oral presentations and poster presentations).
- Making quick and simple approximations to predict experimental outcomes (e.g., back of the envelope calculations).
- Developing expertlike views about the nature of the process of doing experimental physics (e.g., experimentation is iterative, not linear).
- Developing collaboration and teamwork skills.
- Reflecting on and evaluating one's own learning and knowledge (metacognition).
- Enjoying experimental physics and/or the course.

These goals were adapted from several sources, including the American Association of Physics Teachers (AAPT) lab recommendations [6], the EP3 guidelines [70], and previous work [71]. Further, some of these goals also resulted from author brainstorming sessions as well as interviews, as described below.

The goals section went through major revisions during the interview process. First, the Likert scale was changed; initially, it included only major goal, minor goal, and not a goal. However, we observed that many interviewees would say that one of the goals is not a goal of their course, but they would be interested in implementing it. They would then often select "minor goal," despite stating it is not a goal of their course. In order to address this issue, we introduced a fourth Likert option: future goal (not currently a goal). While analyzing the survey data, we currently collapse this category with not a goal, but it helps to provide more accurate results in our data collection.

Additionally, the list of goals presented underwent revisions. Some goals, such as developing communication skills, were split into several goals [in this case, the split was into three goals: developing lab notebook keeping skills, developing scientific writing skills (e.g., lab reports), and developing other communication skills (e.g., oral presentations, poster presentations)]. This was due to interviewee input about the concepts they thought were covered by the goal. In this particular example, when asked to define "communication skills," interviewees had many different ideas about what this might include. Therefore, we split this into three distinct goals in order to collect the most accurate data. Further, we added goals at the request of interviewees, such as "enjoying experimental physics and/or the course" and "making quick and simple approximations to predict experimental outcomes (e.g., back of the envelope calculations)."

Other changes to the goals section as a result of interviews included wording changes to clarify meaning, as well as adding examples to the goals to make them more easily understood.

6. Activities

The first part of this section of the survey asks participants whether they use any officially branded approaches to lab instruction in their course [e.g., Investigative Science Learning Environment (ISLE) Physics [72], Student-Centered Activities for Large Enrollment Undergraduate Programs (SCALE-UP) [73], and Modeling Instruction [74]], as well as whether they use any RBAIs to evaluate the course (e.g., Survey for Physics Reasoning on Uncertainty Concepts in Experiments, or SPRUCE [75,76]; Modeling Assessment for Physics Laboratory Experiments, or MAPLE [77]; and E-CLASS [41]).

The next part of this section lists several activities students might participate in during an undergraduate physics laboratory course and asks how often students engage with them along the following Likert scale: Very frequently, Somewhat frequently, 1-2 times per semester/term, Would like to use in the future, and Never. Details of the development of this Likert scale are given below. The activities probed are divided into the following categories:

- Data analysis and visualization.
- Communication.
- Student decision making.
- Materials and resources.
- Modeling and other activities.

Within the above categories, examples of activities include quantify uncertainty in a measurement, write lab reports, develop their own research questions, and calibrate measurement tools; the full list of activities can be found in Supplemental Material [69].

The activities section also underwent significant changes during the interview process. A question probing whether research-based assessments are used to evaluate the course was added as a result of interviewees mentioning during interviews that they use some of these assessments.

The Likert scale used in the list of activities was changed to make things clearer. Originally, the scale was Always, Often, Sometimes, Rarely, Never. However, this scale was not appropriate for some activities. For example, students might design and present a poster once during a course, and it is unclear which of the scale points this should fall into, because it is not typical for students to make posters "frequently" when compared with other activities, such as quantifying uncertainty in a measurement, which might occur with every lab experiment. Similarly, it is unclear what it means to "always" complete a safety training—some courses might require a single training, whereas others might have a few that students have to complete.

The new Likert scale we implemented is Very frequently, Somewhat frequently, 1-2 times per semester/term, Would like to use in the future, and Never. This scale has several advantages over the previous one. First, it includes an aspirational scale point (i.e., Would like to use in the future), which can be collapsed with Never when analyzing the survey data but again helps discourage those who select a different option despite not using the activity (similar to the aspirational Likert scale point in the goals section). Second, the new scale helps clarify events that might happen only 1 or 2 times in a semester, such as a poster presentation or a safety training. Finally, for activities that might happen more commonly in courses—such as keeping a lab notebook or writing their own code—it provides several scale points that are more easily understood. Because 1-2 times per semester/term is an option, it is clear that "somewhat frequently" means that students participate in the activity more than this, while "very frequently" indicates a higher degree. Therefore, based on interviewee responses to this scale, we feel that we have appropriate knowledge of what it means each time someone selects a particular scale point.

In addition to a new Likert scale, many of the activities were also changed. In some cases, wording was altered or examples were added to clarify meaning. Some activities were combined into one option after it was determined that interviewees could not always tell the difference between them. One example of this is combining "refine experimental apparatus or procedure to reduce random uncertainty" and "refine system to reduce systematic uncertainty" into "refine experimental apparatus or procedure to reduce uncertainty (statistical and/or systematic)". This was due to many interviewees not being able to give appropriate examples differentiating random and systematic uncertainty, and therefore the information obtained from probing these separately was inaccurate.

Finally, some activities were added to this section, such as "engage with PhET simulations" and "complete safety training" due to interviewee requests.

7. Evaluation of students' work

This part of the survey probes how students are graded (evaluated) in the course. The first question probes whether students are assigned individual or group grades, while the second lists potential parts of the course that might factor into student grades and asks participants to select all that are used for their course. Examples include taking a quiz at home before the lab, lab notebooks, lab reports, written exams, poster presentations, practical exams (i.e., hands-on exams), and peer feedback on other students' work. Finally, participants are asked whether they use a rubric (a set of guidelines about how something is graded) to grade student work.

This section of the survey also underwent significant changes due to interviews. Initially, participants were provided with a list of items that might potentially be included in student grades and were asked to rate them on a Likert scale: Not used, Marked/graded for inclusion in final course grade, Marked/graded but not used to determine final course grade. This led to confusion, especially about certain items that are not directly used in the final grade but might be used in some indirect way. For example, attendance at each individual course meeting might make up an overall attendance grade that is then used to determine the final course grade. Interviewees were then uncertain about where on the Likert scale to include attendance. We changed this question to be multiple response and asked participants to select all of the items on the list that are used in grading the course. Because it is a binary option, interviewees understood how to handle indirect effects on grades (they did select these items). Removing the Likert scale helped make the survey more clear.

Additionally, some items were removed (such as reflection questions) due to interviewee lack of understanding around these points, added (such as worksheets) due to interviewee requests, and reworded to help with clarity and understanding.

8. Optional long entry

In this section, two long-form text boxes are provided. Both are listed explicitly as optional; while most of the rest of the survey's questions are optional, these are the only two questions which state this. The first asks participants to enter the titles of lab experiments in any language they wish. This is useful in looking at trends of common laboratory themes that might be present, especially in introductory labs (e.g., during the interviews, many instructors discussed using a pendulum activity in introductory mechanics). While this qualitative data might not provide the most accurate evaluation of themes—some might choose not to include this information, and others might have laboratory titles that do not fully reflect the activities—this will still provide a wealth of information. We encourage any language entry to allow participants to simply copy and paste lab manual titles in order to make this step easier. Online tools such as Google Translate [78] provide an accurate enough translation to qualitatively code themes in later steps of the analysis.

The second text box asks for any additional comments that participants might have. This is useful in cases where the survey may not have fully captured the experience of the participant in teaching their course or for any clarifying comments they would like to make about their prior responses. This further serves to provide content validity, as anything missing from the survey questions themselves can be captured in this box, and therefore, the survey is fully representative of what it aims to measure.

The optional long entry section was not significantly altered during the interview process.

B. Survey results

We present here an overview of physics laboratory courses around the world based on the data we have collected thus far in order to address our second research question. We received responses from 217 unique courses in 41 countries. A figure showing a map of the countries where instructors responded to the survey is shown in

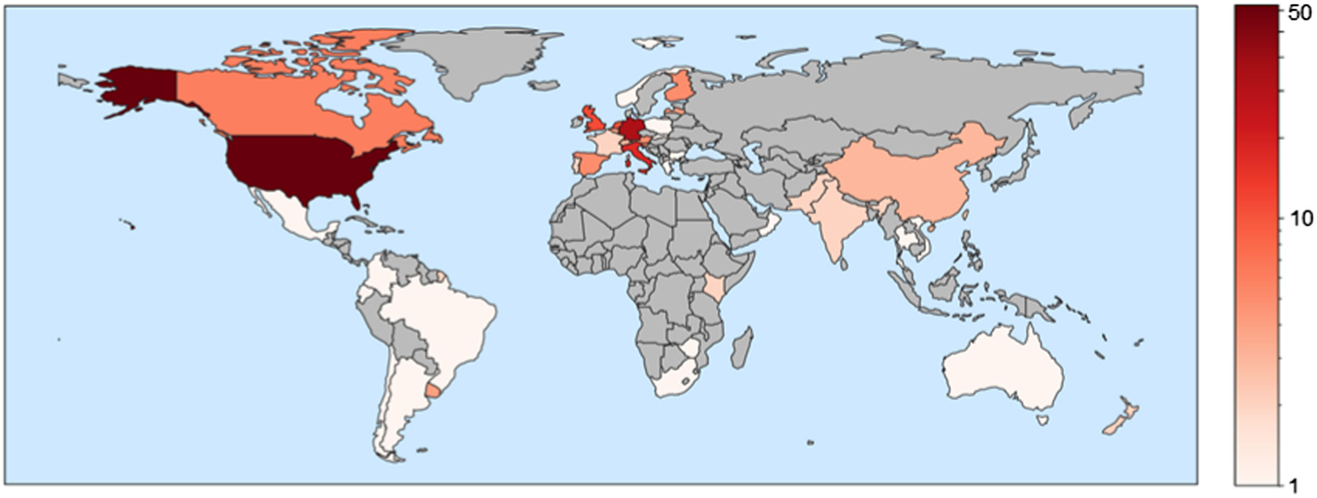

FIG. 2. World map of number of survey responses. Shown on a log scale, each colored country has at least one response; countries in gray have no responses.

Fig. 2, with a full list of the number of respondents per country located in the Appendix.

As discussed previously, our data are skewed toward the authors' countries and is lacking representation in many areas. In future work, we hope to present a more representative sample.

1. Overall characteristics

Of our respondents, most courses (166/217) were offered at Ph.D.-granting institutions, with fewer being offered at Master's-granting institutions (20/217), Bachelor's-granting institutions (25/217) and Associate's-granting institutions (6/217). Most of this variation comes from within the United States—33 of the non-Ph.D. granting institutions are in the United States (63 total responses), whereas only 18 are outside of it (217 total responses).

Of the courses surveyed, 137 are introductory, 79 are beyond introductory, and we have no data about one course. We discuss a split of the data by introductory and beyond introductory where appropriate in our analysis.

Next, we examine the number of students per course and the number of students per section in the course (i.e., the number of students present in the laboratory room at one time). Distributions for both of these are shown in Fig. 3. The median number of students per course is 50. Overall, there are only a few very large courses with more than 500 students. Most courses (188/216) have fewer than 200 students per course. Only three beyond introductory courses have more than 200 students. The median number of students per section is 18. Most courses (152/217) have between 10 and 40 students per section. Very few courses (14/217) have more than 75 students present in the lab at any given time; of these courses, most (12/14) are introductory, with only one beyond introductory.

We probe the topics covered in the course by providing a multiple-response list of possible topics and the option to type in additional topics not listed (no clear patterns emerge from the not listed responses). We present the topics covered in the courses in Fig. 4, with a split shown between introductory and beyond introductory courses. Classical

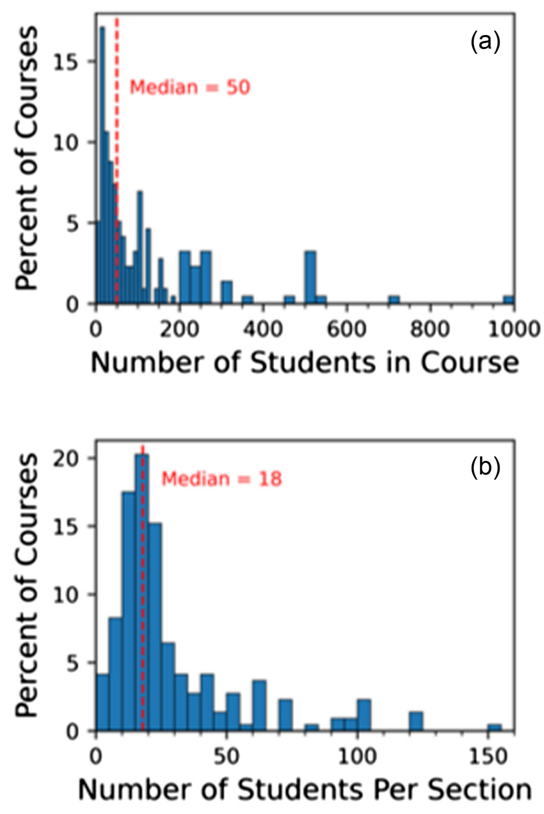

FIG. 3. Number of students per course and per section. Upper histogram (a) shows the distribution of the total number of students in the course, with a median shown as a red dashed line (50) [$N = 216$]. Lower histogram (b) shows the distribution of the number of students per section of the course (i.e., number of students in the lab room at any time), with a median shown as a red dashed line (18) [$N = 217$].

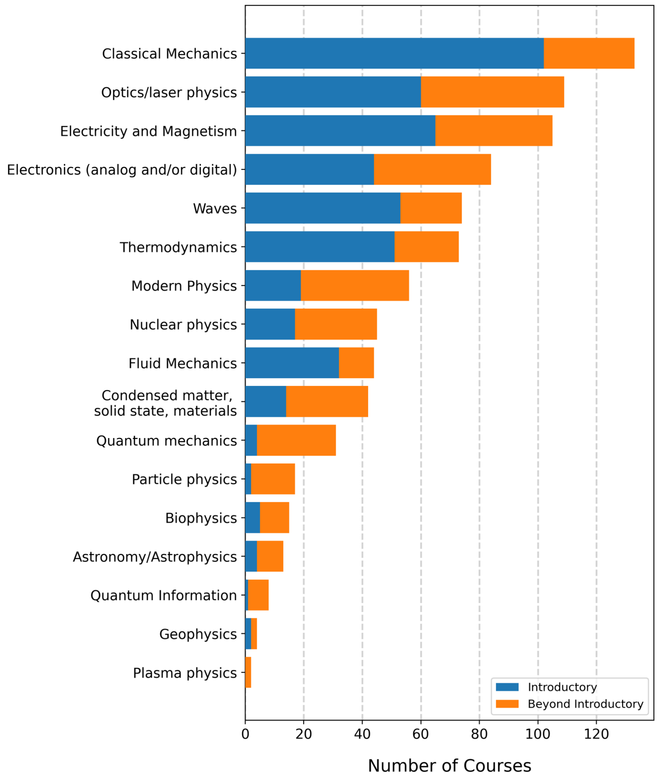

FIG. 4. Topics covered in the course, split by introductory (blue, left) and beyond introductory (orange, right) courses [$N = 216$]. The bottom axis shows the absolute number of courses that included the topic. The most common topic for introductory courses is classical mechanics, and the most common topic for beyond introductory courses is optics and laser physics.

mechanics is the topic selected most often for introductory courses, while optics and laser physics was the most common topic in beyond introductory courses. More specialized areas of physics, such as plasma physics and geophysics, are rarely covered.

In addition to the course topics, we also asked instructors to provide the titles of their lab experiments in a long-form text box. Of the 217 respondents to the survey, 111 chose to do so, and we received 1078 lab titles from these courses. After translating all of the titles to English, we qualitatively coded them to determine the most common types of experiments occurring in undergraduate physics lab courses. These experiments are presented in Table I. This includes codes with at least 15 courses using them. Of the 1078 lab titles received, 96 (8.9%) were uncoded due to being too vague (e.g., classic mechanics), activities beyond a lab experiment (e.g., poster preparation), or too specific (e.g., interaction and collaboration of kirigami). The other 982 lab titles were categorized with one to three codes. The code definitions for all codes and more information about the process of coding the lab titles are located in the Appendix. Additionally, we created a word cloud of these lab titles after removing stop words [79] to give a visual

TABLE I. Most common experiments as given by titles of lab experiments. These are the codes for all categories with at least 15 courses reporting at least one lab in this category. A total of 111 courses providing 1078 lab titles were qualitatively coded to determine the most common experiment types. The definitions of these codes and others not shown are provided in the Appendix.

Topic	Number of courses	Number of lab titles
Optics (intermediate)	68	108
Kinematics	36	55
Dynamics (mechanics)	33	56
Electronics (intermediate)	29	69
Electronics (simple)	29	56
Spectroscopy	28	37
Test and measurement equipment	27	30
Thermodynamics	26	56
Introduction to measurement and uncertainty	25	36
Pendulum	23	33
Optics (simple)	23	39
Particle physics	18	44
Optics (advanced)	17	49
Advanced materials and solid state	17	30
Waves	17	22
Electric fields and electrostatics	16	24
Fluids	15	21
Magnetic fields	15	18

representation of the data (see Fig. 5). We hope that, as we collect more data, qualitative coding of lab titles will help in creating a taxonomy: we can work on grouping courses that complete similar types of experiments.

We also asked whether the laboratory course is integrated with lecture (i.e., if both are one course combined) or if the laboratory course is a separate course. There are

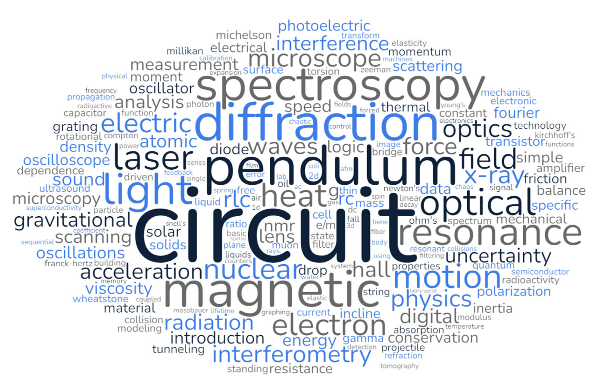

FIG. 5. A word cloud showing the 200 most common words after removing stop words from the lab titles and using basic lemmatization [$N = 111$ courses with 1078 lab titles]. This helps us form a visual representation of the types of experiments happening in undergraduate physics lab courses around the world. Electronics, mechanics, and optics experiments dominate the word cloud.

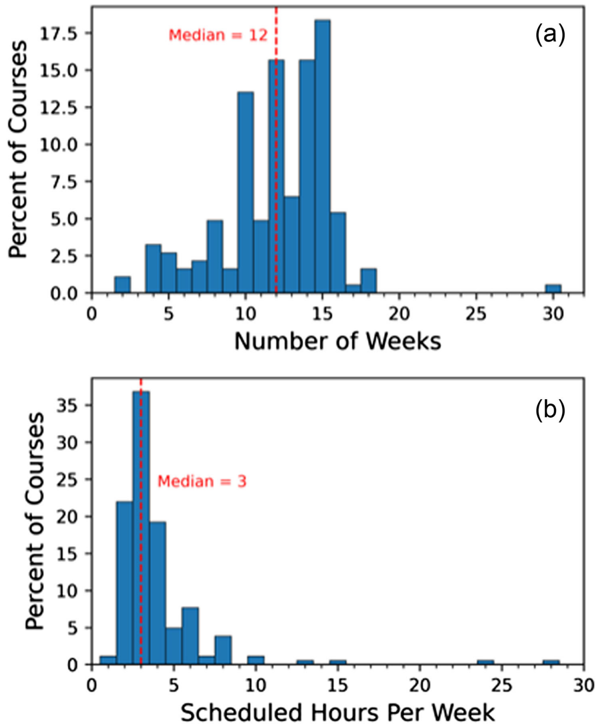

FIG. 6. Number of weeks the course runs for (a) [$N = 185$] and number of hours per week students are scheduled to be in the lab (b) [$N = 182$]. Red dashed lines show the median. The median number of weeks the course runs is 12, and the median number of hours per week is 3.

129 courses that are separate, while 88 are combined with a theory course. Additionally, 136 courses include lectures about statistics and/or data analysis, whereas 81 do not.

Next, respondents reported the number of weeks the course runs for, as well as the number of hours per week students are scheduled to be in the lab. The distributions for these are shown in Fig. 6. The median number of weeks is 12, and the median number of scheduled hours per week is 3. Further, data about the number of hours per week beyond the scheduled time that students spend in the lab are presented in the Appendix; in most cases, students do not spend any time beyond what is scheduled in the lab.

We further investigate the number of experiments per lab meeting students complete and if students have a choice over which experiments they complete. This question is multiple response, and respondents can choose whether students complete multiple experiments per meeting, one experiment per meeting, one experiment per multiple meetings, or a multisession open-ended project. The distribution of responses is presented in Fig. 7. In most courses, students spend time doing one experiment per meeting of the course. Students are often not given a choice of which experiments they do, with 125 courses not allowing students any choice in which experiments they complete, 59 allowing students to choose their experiments

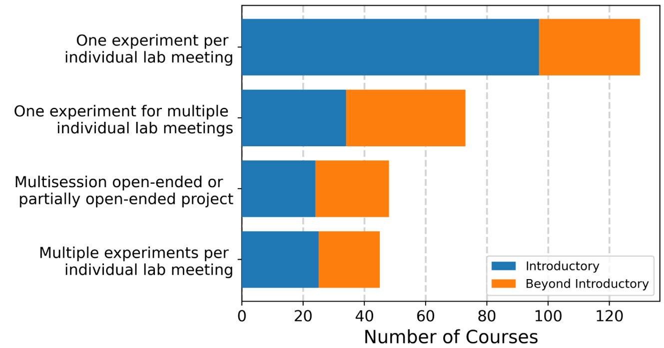

FIG. 7. Types of experiments students complete in the course, split by introductory (blue, left) and beyond introductory (orange, right) [$N = 216$]. This question was multiple response, so respondents could select as many of these options as apply to their course. Most courses involve some component in which students complete one experiment per meeting of the course.

for some portion of the course, and 33 allowing students to choose experiments all of the time.

For those courses where there is a project component, we asked four additional questions about the project. There are 48 courses that contain some project components. Of these courses, students spend a median of 4.5 weeks engaged in project work. In 26 courses, students choose their project topic "all of the time," compared with 16 courses that allow students to choose "some of the time," and six courses that do not allow students to choose their own project topic. About 32 courses allow students to design their own project all of the time, whereas 13 courses allow students to do this only some of the time, and only 3 courses do not allow students to design their own project. Finally, in 26 courses, students always build their own experimental apparatus for the project, while in 19 courses, they do this sometimes and in 3 courses, they never do this.

2. Students

Respondents reported which major(s) their students have in the course, as well as the fraction of physics and astrophysics majors in the course. For the first of these questions, we presented a list of possible majors as a multiple-response question along with the option to write in majors not contained on the list. The second of these questions is multiple choice. We present the results of these items along with a split by introductory and beyond introductory in Table II. Because the question about the major(s) of the students is multiple response, we received many different potential groupings. The table, therefore, represents only the most common groupings (i.e., a minimum of four courses chose this grouping). Nearly, a third of all courses include only physics majors, with 27% of introductory courses and 41% of beyond introductory courses having only physics majors. The second-most

TABLE II. Most common grouping of student majors in a class and percentage of the course that is physics majors. Only the most common groupings are shown (i.e., a minimum of four courses in that grouping).

	% Responses ($N = 217$)	% Responses, Intro ($N = 137$)	% Responses, Beyond Intro ($N = 79$)
Physics	32.2	27.0	41.8
Physics, astrophysics/astronomy	9.2	5.1	15.2
Another science (e.g., biology and geology)	5.5	8.8	0.0
Physics, physics/astronomy teaching/pedagogy	5.5	2.2	11.4
Engineering	4.6	5.8	2.5
Physics, engineering	2.8	14.6	5.1
Chemistry, another science (e.g., biology and geology)	2.3	3.6	0.0
Physics/astronomy teaching/pedagogy	1.8	2.2	1.3
Other	35.9	43.8	22.3

	% Responses ($N = 216$)	% Responses, Intro ($N = 136$)	% Responses, Beyond intro ($N = 79$)
0%–25% physics majors	34.3	50.0	7.6
25%–50% physics majors	7.9	6.6	10.1
50%–75% physics majors	2.8	8.1	6.3
75%–100% physics majors	50.5	35.3	75.9

common combination overall is physics and astrophysics or astronomy majors, which accounts for another 9% of responses. Overall, 166 out of the 217 courses surveyed (76%) included physics majors, and therefore, 51 courses (24%) do not include any physics majors.

Another point of discussion about majors is that the United States typically treats physics courses differently than courses outside of the United States. Within the United States, it is very common to combine many different majors into one course at the introductory level, whereas outside of the United States, it is more common to have an introductory physics course only for physics majors, one for those training to teach high-school physics, a separate course only for engineering majors, etc. When examining only introductory courses, 81% (29 out of 36) introductory courses in the United States contain 0%–25% physics majors while outside of the United States, this number drops to 39% (39 out of 101 courses).

Next, we provide information about how many years the students have been at the University when they take the course. Again, this question is multiple response, so instructors can choose as many options as apply to their course. These data are presented in Table III. The courses in our dataset lean heavily toward first-year and second-year students.

TABLE III. Year of students in the course [$N = 216$]. This question is multiple response, so instructors can select all options that apply to their course.

	Number of responses	% Responses
1st year	106	49.1
2nd year	83	38.4
3rd year	66	30.6
4th year	39	18.1
5th year or higher	9	4.2

3. Students working in groups

We next inquire about the ways in which students work together in the course. Our survey results showed that 204 courses indicated students work with at least one partner, and 13 courses indicated that students work alone. Data about these 204 courses are shown in Table IV, including the number of lab partners, whether students stay in the same group for the entire course, and whether students choose their own groups. In most cases, students are working in pairs of their own choice and stay with this lab partner for the entire term.

TABLE IV. Grouping of students [$N = 204$]. Most students work with one lab partner of their choice for the entire term.

	Number of responses	Percent responses
Groups of 2	118	57.8
Groups of 3	63	30.9
Groups of 4	19	9.3
Groups of 5+	9	4.4
Stay in same groups	165	80.9
Switch groups	39	19.1
Choose their groups	139	68.1
Are assigned groups	31	15.2
Both options	34	16.7

TABLE V. Mean number of students per staff member in the lab. We pair the question probing number of staff in the room with the question about the number of students in each section to determine these averages. On average, there are a total of 9.9 students per staff member. Means take into account only courses with at least one of that type of instructional staff (e.g., courses with no undergraduate LAs are not counted in the mean number of students per LA).

	Mean	Number of courses
Students per faculty	25	186
Students per lab technician	33	95
Students per graduate TA	20	116
Students per undergraduate LA	21	51
Students per staff (total)	9.9	215

4. Instructional staff

This section of the survey asks about the number of different types of instructional staff present in the lab room with the students. Because we also know the number of students present in the lab at one time, we can determine the average number of students per staff. The means of this are shown in Table V, including information about faculty members, lab technician, graduate and postdoctoral TAs, and undergraduate LAs. The distribution of the number of students per staff member is shown in Fig. 8. The mean number of students per staff member (after summing all possible types of staff members) is 9.9. Few courses utilize LAs, whereas many courses have faculty and TAs present with students.

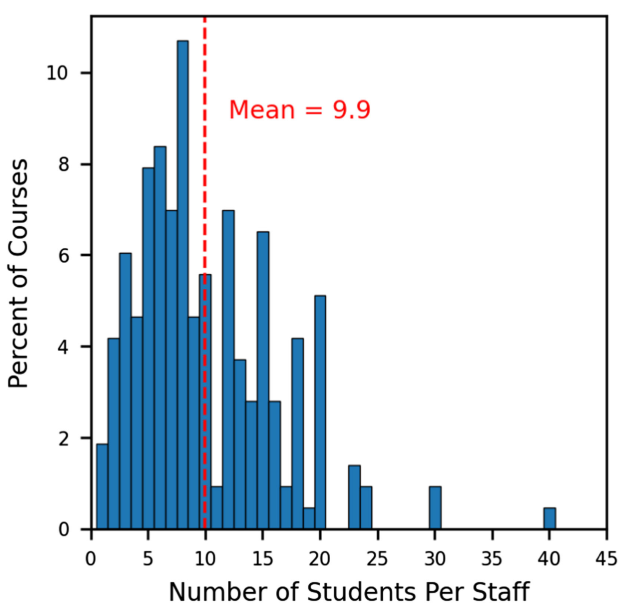

FIG. 8. Distribution of the number of students per staff member in the lab room at any given time, with mean = 9.9 shown as a red dashed line.

Further, respondents provided information about the frequency and type of training for both graduate TAs and undergraduate LAs. This question is multiple choice in which respondents can indicate whether training happens once per term, once per academic year, or weekly. The types of training is a multiple response question, which allows respondents to select pedagogy, grading, and familiarization with lab equipment in any combination that applies to their course. Both of these questions also have "not listed" options with the opportunity to write in a response; no patterns emerged from an analysis of these responses. These data are shown in Table VI. Nearly all courses that train TAs and/or LAs provide instruction to familiarize them with the lab equipment, while more than half also offer pedagogy or grading training. There is no standardized frequency of this training, with about one-third providing training once per term and one-quarter providing training once per academic year or weekly.

5. Goals, activities, and evaluation

Potential course goals or learning objectives were presented as a list with options to select "Major Goal," "Minor Goal," "Not a Goal," and "Future Goal (not currently a goal)." As previously discussed, the latter two of these categories are collapsed for all analyses. Each of the 15 goals presented had between 215 and 217 responses. A plot of the answers to this question is shown in Fig. 9. Other course goals are possible, but there is no "not listed" option available for this question. The current list of goals was refined through interviews, including the addition of extra goals as requested by interviewees.

We also examine the total number of goals (major plus minor) selected for each course. This distribution is shown in Fig. 10 and the mean is 11.8 goals out of the possible 15. On average, courses have 6.9 major goals and 4.9 minor goals. There are no significant differences in the number of

TABLE VI. TA and LA training frequency [$N = 137$] and type [$N = 136$]. For the frequency question, respondents can type in an answer if none of the provided options capture their training schedule. Type of training is a multiple-response question and also includes the ability to type in a response if a type of training is missing from the provided options.

	Number of responses	Percent responses
Once per term/semester	45	32.8
Once per academic year	35	25.5
Weekly	33	24.1
Other	22	16.1
Familiarization with equipment	130	95.6
Pedagogy	80	58.8
Grading	80	58.8
Other	7	5.1

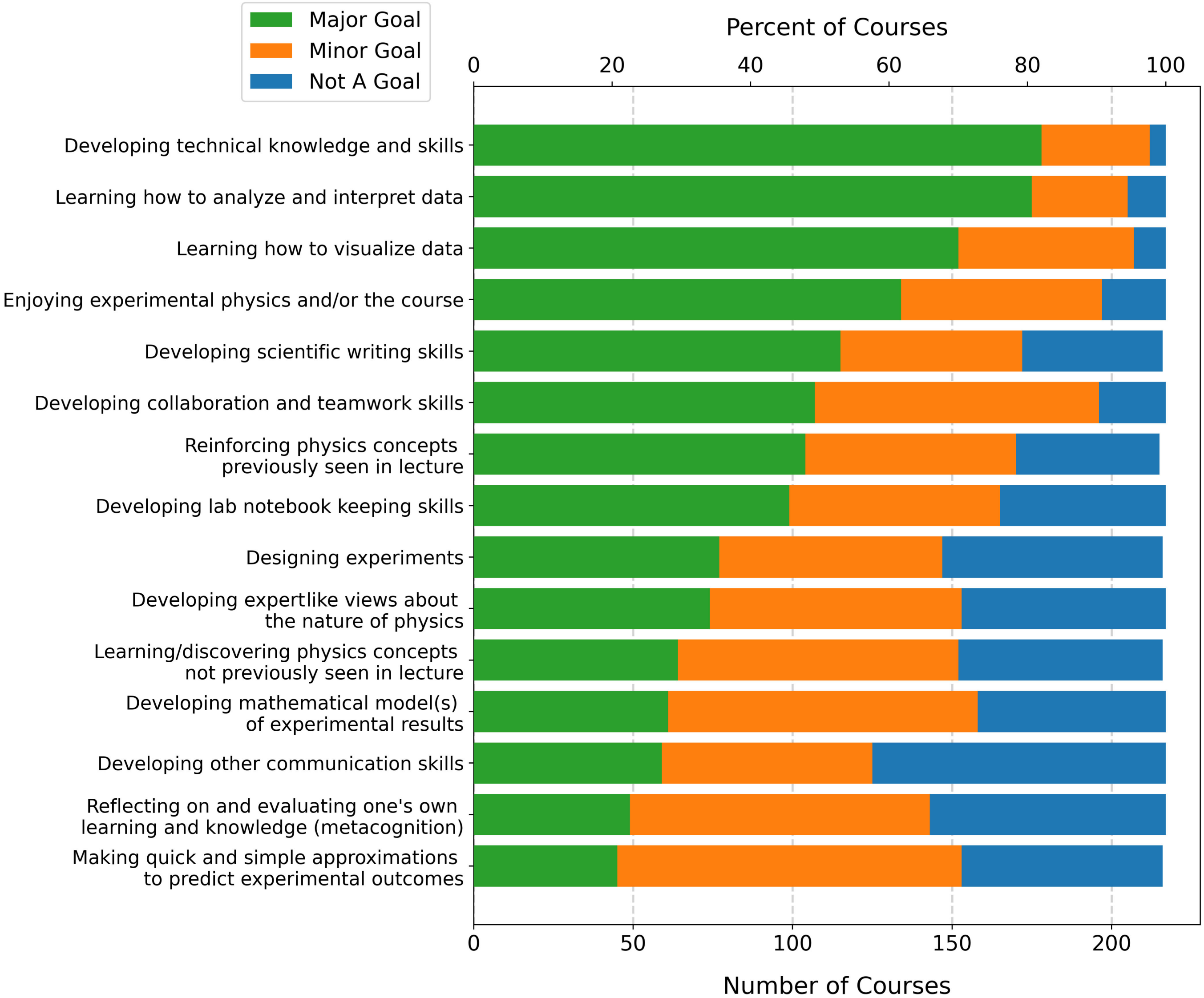

FIG. 9. Course goals, split by major goal (green, left), minor goal (orange, middle), and not a goal (blue, right) for the course. Between 215 and 217 courses provided data for each goal, and the percentages are calculated using the full 217 courses for display purposes. The most commonly selected goal is developing technical knowledge and skills.

course goals (either total, major, or minor) for introductory and beyond introductory courses.

The first question in the activities section asks about whether a specific branded instruction technique is used (such as modeling instruction, SCALE-UP or ISLE). Most courses (135/212) do not use any type of branded instruction method. Similarly, a question about RBAIs reveals that most courses (148/212) do not use any of these.

Next, we present respondents with a list of 41 possible activities broken up into five categories—data analysis, communication, student decision-making, materials, and modeling/other activities. Plots of the responses to the Likert-style questions for each of these categories are shown in Figs. 11–15. These figures are broken down by Likert response (very frequently, somewhat frequently, 1–2 times per semester/term, and never, where we have again collapsed an aspirational scale point with never). We find that courses engage in a wide variety of activities. In some cases, such as in the collection of activities relating to both data analysis and student decision making, at least half of the courses selected that they participated in all activities to some extent. The split between the frequencies students engage in activities also generally occurs as expected. For example, students typically write lab reports very frequently, but design and present a poster 1–2 times per term. The student decision-making category, in particular, has a large number of activities with responses of 1–2 times per term.

Further, we find that students rarely use a scientific paper to guide their lab experiments—in most cases, they either

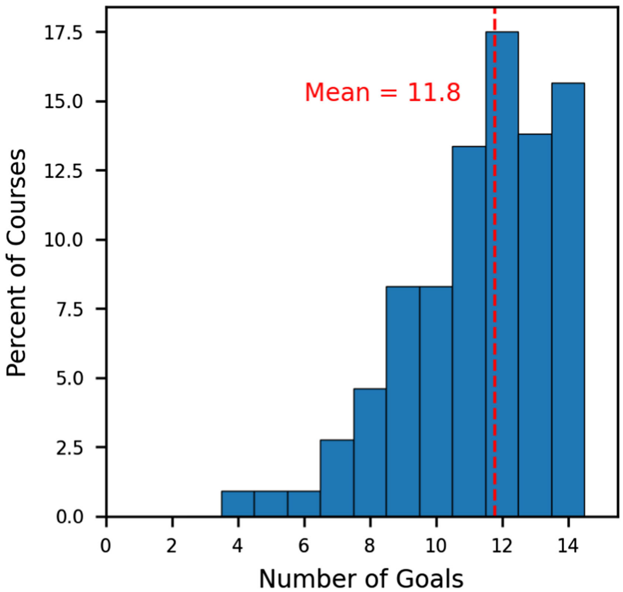

FIG. 10. Distribution of the total number of goals (major plus minor) for each course. The maximum possible number of goals is 15 (all of the provided course goals). The mean number of goals per course (red dashed line) is 11.8.

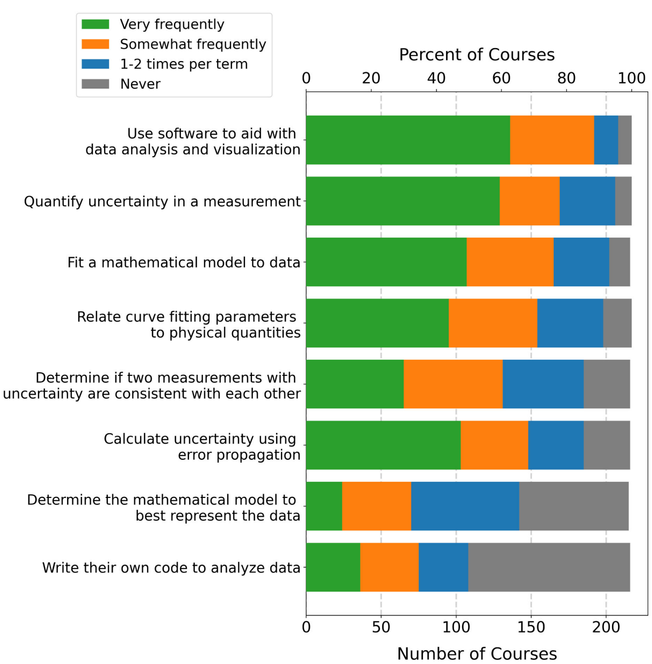

FIG. 11. Courses with data analysis activities [$N = 215$–217]. Bars represent number of courses (bottom axis) and percent of courses (top axis) that include various activities related to analyzing data, such as error propagation and curve fitting. Bars are split based on frequency of the activity—very frequently (left, green), somewhat frequently (second from left, orange), 1–2 times per term (second from right, blue), and never (right, gray). In general, students participate in each of the data analysis activities to some extent in at least half of all courses. The least popular activity was for students to write their own code to analyze data.

use a step-by-step lab manual or a semiguided lab manual. In more than 80% of courses, students use a preconstructed apparatus to some extent. Further, in most courses (more than 80%), students spend some time determining results already known to the instructor, but not yet known to the students, though in nearly 80% of courses, students engage with activities where they are confirming results they have already learned in a lecture course.

A list of potential items that might be graded for inclusion in a student's final course grade is presented to respondents, and they are able to select all of the ones they use in their own course to assign student grades (multiple response). This list of 18 potential things might not fully span the space of items included in a grade, and so

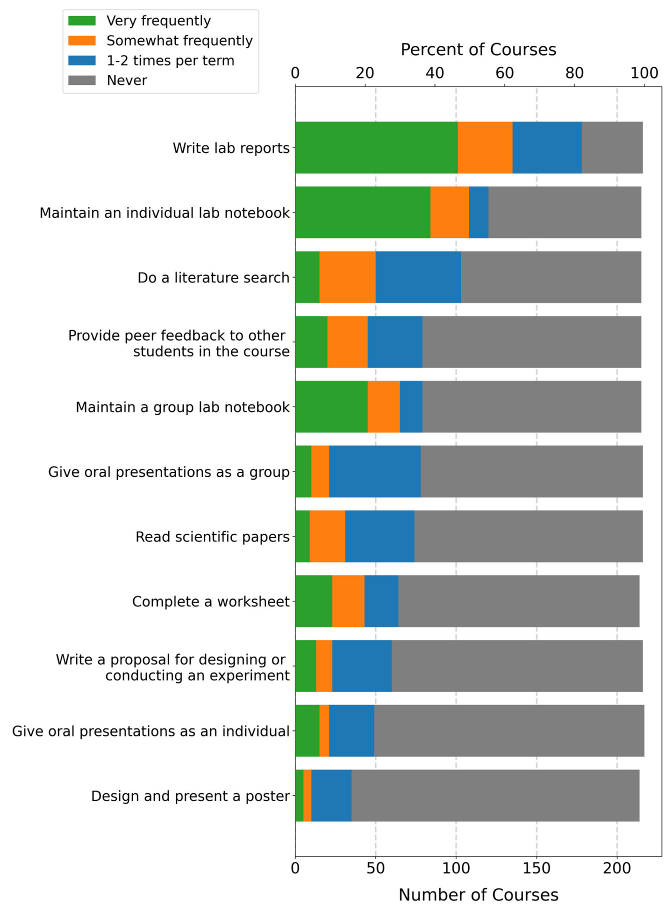

FIG. 12. Courses with communication activities [$N = 214$–217]. Bars represent number of courses (bottom axis) and percent of courses (top axis) that include various activities related to communication skills, such as peer feedback and lab reports. Bars are split based on frequency of the activity—very frequently (left, green), somewhat frequently (second from left, orange), 1–2 times per term (second from right, blue), and never (right, gray). Writing lab reports is a common activity, with more than 80% of courses indicating that students engage in this to some extent. Very few courses have students present posters, give presentations, or write proposals for experiments.

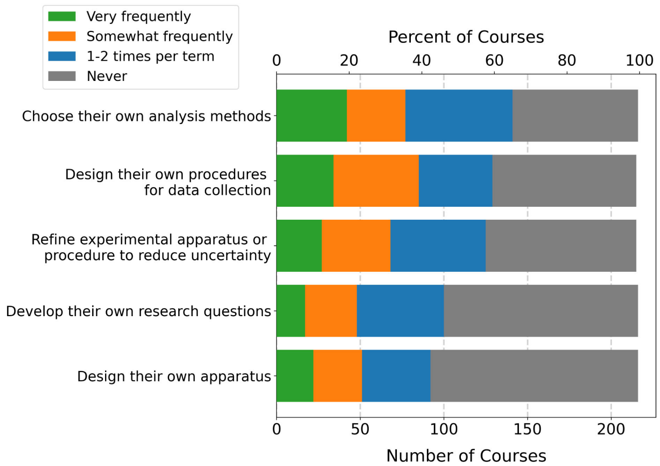

FIG. 13. Courses with student decision-making activities [$N = 215$–217]. Bars represent number of courses (bottom axis) and percent of courses (top axis) that include various activities related to decisions made by students, including choosing their own procedures and analysis methods. Bars are split based on the frequency of the activity—very frequently (left, green), somewhat frequently (second from left, orange), 1-2 times per term (second from right, blue), and never (right, gray). Students engage in these activities in many courses, although in most cases, they are only doing these 1-2 times per term as opposed to other activity categories that have more responses in the very frequently category.

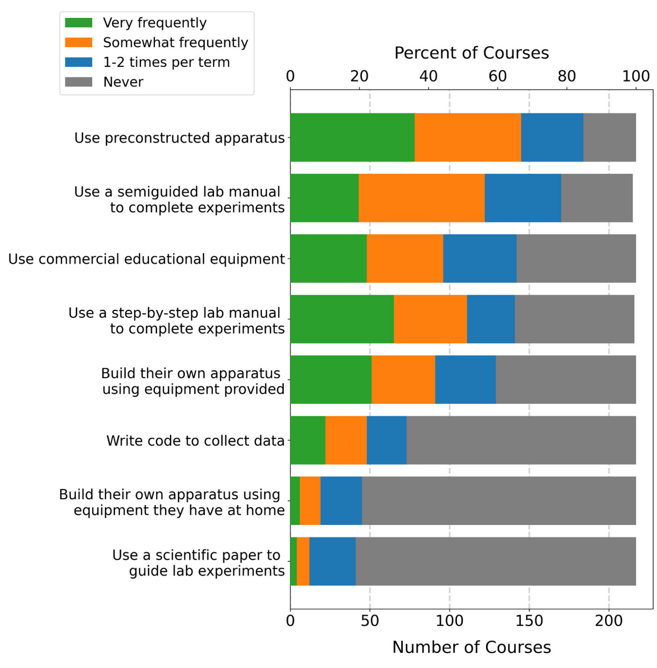

FIG. 14. Students' engagement with materials in lab courses [$N = 215$–217]. Bars represent number of courses (bottom axis) and percent of courses (top axis) that include various methods of student interaction with materials, including use of commercial equipment (such as PASCO or TeachSpin), as well as students building their own apparatus. Bars are split based on the frequency of engagement with the activity—very frequently (left, green), somewhat frequently (second from left, orange), 1-2 times per term (second from right, blue), and never (right, gray).

"not listed" with an option to write in other items is included in the survey. Table VII shows the number of courses that include each option. The analysis of "not listed" write-in responses did not reveal any patterns. Most courses (about 75%) use lab reports to assign grades to students, with attendance and participation being the second most common item with more than half of the courses using this.

Finally, we can combine goals, activities, and items graded to determine how instructors are attempting to meet their course goals as discussed previously. We present the results of this analysis in Table VIII. In this analysis, we collapse the categories of major goal and minor goal together.

We find that courses typically engage in a higher percentage of activities related to a goal than items graded related to that goal; this percentage is often much larger (in some cases, more than 4 times). There is no correlation between having a certain goal for the course and the percentage of activities related to that goal or the percentage of items graded related to that goal, thus showing that instructors take many different paths in attempting to achieve their course goals.

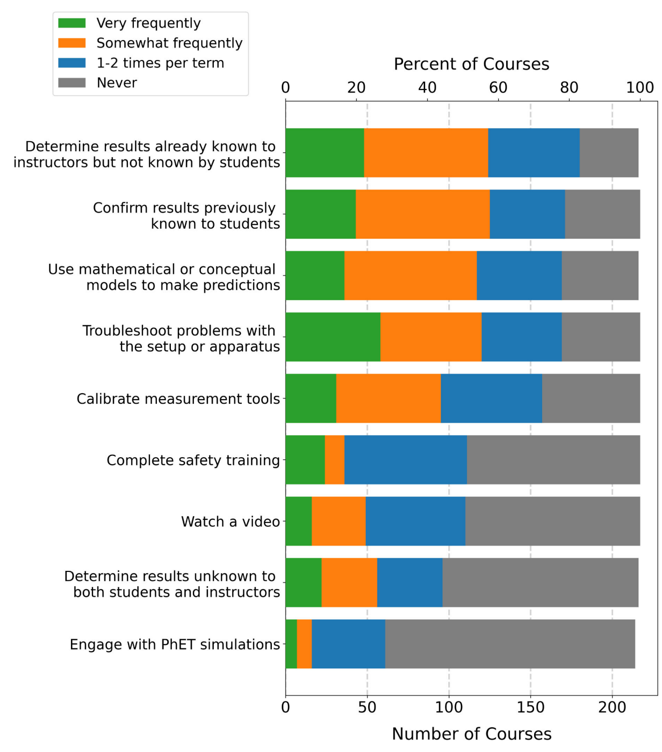

FIG. 15. Courses with modeling and other activities [$N = 214$–217]. Bars represent number of courses (bottom axis) and percent of courses (top axis) that include various activities related to modeling, such as using models to make predictions, as well as other activities such as watching a video or completing a safety training. Bars are split based on the frequency of engagement with the activity—very frequently (left, green), somewhat frequently (second from left, orange), 1-2 times per term (second from right, blue), and never (right, gray).

TABLE VII. Items included in final course grade [$N = 216$]. This multiple-response item allows respondents to select multiple items, as well as an option to write in anything not listed.

	Num. responses	Percent responses
Lab report	160	74.1
Attendance/participation	117	54.2
Lab notebooks	99	45.8
Accuracy/precision of results	88	40.7
Oral presentation	69	31.9
Observation of students	62	28.7
Prelab calculations	59	27.3
Written exam	51	23.6
Interview/meeting after the lab	40	18.5
Worksheets	33	15.3
Partial lab report	31	14.4
Prelab measurement/analysis plan	29	13.4
Practical exam	27	12.5
Quiz/interview prior to working	24	11.1
Poster presentation	22	10.2
Prelab quiz	22	10.2
Peer feedback	19	8.8
Prelab video	18	8.3
Not listed	31	14.4

TABLE VIII. Matching of goals with activities and items graded. The number of courses represents those who selected the goal as either a major or minor goal for their course. This table shows, if an instructor selects a goal, how many activities and items graded they have selected on average (mean) that match that goal. These are shown as fractions as well as percentages. The fractions allow visualization of the total number of matched activities and items graded for each goal, while the percentages allow for comparison between these matched items more easily. One goal (enjoyment of experimental physics and/or the course) is not shown in this table because no activities or items graded match that goal. Additionally, the goal related to approximations has no items graded matched with it (though it does have matched activities). In general, we find that instructors have a higher percentage of activities for a specific goal than items graded for that goal.

	Number of courses	Activities		Items graded	
		Fraction	Percent	Fraction	Percent
Developing mathematical model(s) of experimental results	158	5.7/7	81	1.8/5	36
Making quick and simple approximations to predict experimental outcomes	153	1.6/2	80	N/A	N/A
Learning how to analyze and interpret data	205	8.4/11	76	1.6/5	32
Reinforcing physics concepts previously seen in lecture	170	7.0/10	70	2.8/10	28
Developing scientific writing skills	172	1.3/2	65	1.2/3	40
Learning physics concepts not previously seen in lecture	152	9.1/14	65	2.2/6	37
Learning how to visualize data	207	2.5/4	63	1.4/4	35
Developing expertlike views about the nature of the process of doing experimental physics	153	18.2/30	61	2.4/8	30
Designing experiments	147	5.8/10	58	0.59/2	30
Developing lab notebook keeping skills	165	1.1/2	55	0.57/1	57
Developing technical knowledge and skills	212	3.2/6	53	1.4/4	35
Reflecting on and evaluating one’s own learning (metacognition)	143	0.44/1	44	0.10/1	10
Developing collaboration and teamwork skills	196	1.2/3	40	0.66/2	33
Developing other communication skills	125	2.0/5	40	1.1/5	22

V. CONCLUSIONS AND FUTURE RESEARCH

We have presented the development of a survey designed to collect data that will allow us to create a taxonomy of lab courses with additional data collection. The goals of this paper were to detail the development and validation of this survey, as well as to present initial findings from the data collected.

We detailed the steps of developing the survey, including the initial brainstorming sessions, collection of information according to Van den Akker's Spinnenweb categories [55], and organization of the questions into the final version of the survey. Further, we performed interviews with lab instructors in order to determine that we have evidence of validity of the survey. We have presented evidence for construct, content, and face validity through the development of the survey, both in the initial stages of constructing the items and via the interviews.

We have established content validity through use of the SpinnenWeb framework and the expertise of the authors to construct the initial survey. Further, the initial results from the survey demonstrate that the survey is able to capture a wide range of different courses, which is also related to content validity. We have shown content validity in that the interviewees agreed that the items and response options cover all aspects of a lab course. Finally, we determined face validity in the interviews by showing that the wording of the questions was unambiguous.

We also collected and analyzed survey responses from 217 courses in 41 countries, to present an initial view of physics lab courses worldwide. This global study found many similarities between these courses—for example, in almost all courses, students work with at least one partner. Additionally, a goal of nearly all courses is for students to develop technical knowledge and skills. We also found many differences between these courses, such as the number and types of goals of the course, the activities students participate in, and the student-to-staff ratio. Further, we determined that there is no significant difference between introductory and beyond introductory courses in terms of the number of course goals they have. We also showed that in terms of data analysis and student decision making, at least half of all courses participated to some extent in all activities in these categories.

We also presented interesting results about the level of guidance provided to students, especially that students are much more likely to use preconstructed apparatuses (rather than building their own) and are often engaging in activities that confirm results already learned in a lecture course. More research on open-ended lab courses and how they differ from traditional courses could be useful for understanding the spectrum of guidance students receive in lab courses.

The results of this study have also raised several new questions in investigating undergraduate lab courses. For example, we find that students tend to work with other students in these courses, but we do not explicitly know why courses are structured in this way. It might be due to equipment limitations, logistical constraints due to two or more people being necessary to actually perform the experiment, pedagogical reasons, or perhaps simply tradition. We frequently assume that working collaboratively has pedagogical benefits, but this study does not investigate those, though group work in lab courses is addressed in other PER literature [80–84]. Further, we have no information about how well instructors are meeting their course goals. Because the average number of goals per course is so high (mean $= 11.8$), it seems unlikely that instructors can focus equal time to all of these goals. Even considering major goals only, courses have a mean of 6.9 major goals, which is a significant number of goals for a short course that might only meet a few hours per week for a term. We have no detailed information about actions instructors take to meet these goals aside from knowing some of the activities and graded items that might relate to these goals. Thus, the initial results of this survey suggest many ideas for future PER in lab courses.

We hope to continue data collection to make more claims based on an expanded dataset and eventually build a taxonomy of laboratory courses to help classify these courses and make comparisons easier. We need at least an order of magnitude more data to accomplish this goal, as we would eventually like to use a clustering analysis to analyze the data and find clusters that represent different types of courses [18,85].

Further, if we collect a significant amount of data in individual countries, we would like to analyze the landscape of undergraduate physics laboratory courses in these specific countries; this would require many more responses from each individual country.

Finally, with more data, we can present a more complete view of the landscape of undergraduate physics lab courses worldwide to give instructors and researchers a broad perspective as they work to improve physics laboratory instruction globally.

ACKNOWLEDGMENTS

The authors thank Rachael Merritt for her contributions in helping match goals, activities, and items graded. We also would like to thank the 23 instructors who participated in interviews for the generosity of their time and invaluable help in developing our survey, as well as the course instructors who filled out the survey once it was finished. We would like to thank the Imperial College London European Partners Fund which supported the workshop where the need for this work was first identified. This work is supported by NSF DUE 1914840, DUE 1913698, and PHY 2317149. Publication of this article was funded by the University of Colorado Boulder Libraries Open Access Fund.

APPENDIX A: AUTHOR PROFESSIONAL POSITIONALITY STATEMENTS

As the work presented here relies, in part, on the professional experience and networks of the authors, we present our backgrounds relevant to the work. Further, the backgrounds of the authors are relevant in our first steps toward face, content, and construct validation.

Author G. G. is a Ph.D. student currently working in PER on the development, validation, and analysis of a measurement uncertainty-focused RBAI for use in undergraduate physics laboratory courses. She has prior experience both as an undergraduate and as a graduate student working in experimental physics; her undergraduate career focused mainly on experimental atomic, molecular, and optical physics while her graduate experimental physics experience is in biophysics. Finally, she has been a teaching assistant in a sophomore-level undergraduate physics lab course at the University of Colorado Boulder. She was not present at the workshop where the initial idea for the taxonomy work originated but was brought in sometime later to work on the project.

Author M. A. is the laboratory course coordinator at the University of Potsdam, which she completely redesigned using research-based results. Her physics background is in solid-state physics, but she is now focusing on PER. During her Ph.D., she built a low-temperature scanning tunneling microscope and used it to investigate the manipulation of isolated molecules on surfaces. During her postdoc, she studied the electrical transport properties of graphene. She worked as an experimental physicist both in United Sates and Germany and taught lab courses in both these countries. As a student, she did her lab courses in Italy. She is an active member of the working group about undergraduate laboratory courses called Arbeitsgruppe Physikalische Praktika [86] of the German Physical Society (DPG) and since 2021 served as a board member for the physics education section of the DPG.

Author M. F. J. F. has been the second-year laboratory course coordinator at Imperial College London since 2023. His research is focused on student learning in undergraduate teaching laboratories and equity in physics. He completed a Ph.D. in theoretical plasma physics, taught high-school physics in London for 3 years, and worked as a postdoctoral researcher with H. J. L. in PER related to laboratory courses completing work on the quantum industry [87], as well as analysis of E-CLASS data [53] and development of the MAPLE survey [77].

Author P. S. W. M. L. is the laboratory course coordinator at the Leiden Institute of Physics in The Netherlands. His physics background is in applied physics, but since 2009, he is focusing on physics education. From 1993 to 2014, he worked as a high school physics teacher. In 2014, he finished his Ph.D. in physics education in which he used educational design research to develop a teaching-learning sequence on the general law of energy conservation in which students reinvent that law by performing various experiments. After finalizing his Ph.D., he started to work in Leiden where he has been redesigning the undergraduate lab courses since 2017 using educational design research. He is a board member of the Groupe International de Recherche sur l'Enseignement de la Physique (GIREP—International Research Group on Physics Teaching) [88] and coleader of the GIREP Thematic Group on Laboratory Based Teaching in Physics (LabTiP) [89].

Author E. T. is a Ph.D. student actively involved in PER. His research focuses on the integration of active learning methods in both introductory university physics courses and high school settings (in particular, using the ISLE approach). With many years of experience as a high school physics teacher, he has developed a keen interest in the use of digital technologies to enhance physics learning. He has been involved in the introduction of the E-CLASS assessment tool in Italy. E. T. attended the workshop in April 2022, where he presented the first results of the implementation of E-CLASS in Italy.

Author H. J. L. is a physics professor who runs two research groups, one in PER and one in experimental chemical physics. Her Ph.D. was in the field of experimental atomic physics (Bose-Einstein Condensation), and her postdoc was in experimental molecular physics (Cold molecule spectroscopy). She has over 25 years of experience designing, constructing, and using tabletop experimental apparatus. Her work in PER began in 2010 and has focused mostly on laboratory courses at the undergraduate level. Besides the current work, she has had many international collaborations in PER, including with researchers from China, Oman, South Africa, Denmark, Norway, United Kingdom, Italy, and Germany. Additionally, she has taught all of the undergraduate physics laboratory courses at the University of Colorado numerous times. She also led efforts to transform three of these courses through research-based practices. She has served on the Board of Directors of the Advanced Laboratory Physics Association (ALPhA) [64] for 11 years, including 2 years as president of the organization. She was present at the workshop in April 2022 that initiated this project.

APPENDIX B: ADDITIONAL SURVEY RESULTS

In this Appendix, we present a list of countries represented in the interviews and additional results from the survey.

Interviews took place with instructors from the following countries: Australia, Brazil, Canada, Chile, China, Colombia, Finland, France, Georgia, Greece,

TABLE IX. Respondents by country. The number of unique courses per country that are included in the final dataset ($N = 217$), as well as the percent of responses from each country are listed. The countries where the authors are from have a higher than average percentage of responses.

Country	Number of responses	% Responses
United States	63	29
Germany	33	15
Italy	17	7.8
United Kingdom	11	5.1
Netherlands	10	4.6
Canada	6	2.8
Hong Kong	6	2.8
Austria	5	2.3
Finland	5	2.3
Slovenia	5	2.3
Spain	5	2.3
Belgium	4	1.8
Latvia	4	1.8
Uruguay	4	1.8
China	3	1.4
Switzerland	3	1.4
Czech Republic	2	0.92
France	2	0.92
Kenya	2	0.92
India	2	0.92
New Zealand	2	0.92
Pakistan	2	0.92
Portugal	2	0.92
Taiwan	2	0.92
Argentina	1	0.46
Australia	1	0.46
Brazil	1	0.46
Bulgaria	1	0.46
Chile	1	0.46
Colombia	1	0.46
Ecuador	1	0.46
Greece	1	0.46
Mexico	1	0.46
Norway	1	0.46
Oman	1	0.46
Poland	1	0.46
Slovakia	1	0.46
South Africa	1	0.46
Thailand	1	0.46
Vietnam	1	0.46
Zimbabwe	1	0.46

TABLE X. Number of hours per week beyond the scheduled time students spend in the lab [$N = 185$]. In most courses, students do not spend any time in the lab other than what is scheduled.

	Number of courses	% Responses
0 h	103	56
1–3 h	56	30
4–6 h	7	3.8
More than 6 h	6	3.2
Unknown	4	2.2
Not allowed	9	4.9

TABLE XI. Definitions of codes used to qualitatively code the lab titles as well as the number of courses and number of lab titles coded for each. We include in some cases specifically items that are not included as part of the code.

Code	Definitions	Number of courses	Number of lab titles
Advanced materials and solid state	Crystals (including 2D crystals), ferrite hysteresis, fluorophore characterization, magnetic hysteresis, nanoparticles, photovoltaics, plasmon resonance, PN junctions, quantized conduction, quantum dots, quantum Hall effect, semiconductors, solar power, superconducting quantum interference device (SQUID), superconductivity, surface physics, surface roughness via advanced microscopy techniques, thermionic emission, tribology	17	30
Arduino		3	3
Blackbody radiation	Blackbody radiation, Planck radiation, thermal radiation	5	5
Charge-to-mass ratio of electron		8	8
Density measurement	Archimedes' principle, measuring the density of liquids and/or solids	8	9
Dynamics (mechanics)	Atwood machine, collisions, conservation laws (energy, momentum), energy, drag, forces, friction, Maxwell wheel, measuring gravitational constant G (NOT measuring gravitational acceleration g), Newton's laws, orbits, torque, work; NOT pendulum, NOT springs	33	56
Electric fields/electrostatics	2D electric potential [V(x,y)], capacitance, Coulomb's law, current balance, dielectric properties of materials, forces between capacitor plates	16	24
Electron diffraction		6	6
Electronics (advanced)	Adders, central processor creation, counters, decoders, digital circuits, digital microchips, drivers for hardware devices, electronic oscillator, flip-flops, Fourier transform, harmonic oscillator circuit implementation, logic gates, multiplexors, Nyquist-Shannon sampling theorem, registers, serial adders, small radio construction, stopwatch with OLED display, triodes	10	30
Electronics (intermediate)	Chaotic circuits, coaxial transmission line, diodes, electric engines, electronic feedback and/or control, electronic hysteresis, impedance, IV characteristics (NOT Ohm's law), light-emitting diodes (LEDs), lock-in amplifier, magnetoresistive effects, motors, nonlinear circuits, operational amplifiers (op-amps), passive filters, power factor measurement, rf spectrum analyzer, RLC circuits, Thevenin circuits, toggle circuits, torsion magnetometer, transients, transistors, Wheatstone bridge	29	69
Electronics (simple)	ac circuits, capacitors, dc circuits, inductors, internal resistance, Kirchoff's laws, material resistivity and/or wire resistivity, Ohm's law, resistors, RC circuits, RL circuits, series and parallel circuits, voltage sources	29	56
Fluids	Aerodynamics, Bernoulli's equation, Brownian motion, diffusion, fluid flow, Hagen-Poiseuille's law, liquids, rheological behavior, stable Kaye effect, Stokes law, superfluid helium, surface tension	15	21
Franck-Hertz experiment		6	6
Hall effect		10	11
Interferometry	Fabry-Perot, Mach-Zehnder, Michelson	12	14
Introduction to measurement and uncertainty	Generic "introduction to equipment" or "introduction to measurement,"control of variables, statistics lectures, uncertainty analysis and/or error propagation	25	36
Kinematics	Center of mass, free fall, gyroscope, inclined plane, measuring gravitational acceleration, g (NOT gravitational constant, G, and NOT the use of a pendulum), moment of inertia, motion analysis, parabolic motion, projectile motion, rotational motion	36	55
Lasers	Diode laser, fiber laser, laser pulses, Nd:YAG laser; NOT HeNe lasers, NOT laser spectroscopy	11	12
Magnetic fields	Earth's magnetic field, eddy currents, Faraday's law and/or induction, Helmholtz coils, induced electromotive force, solenoids	15	18

(Table continued)

TABLE XI. *(Continued)*

Code	Definitions	Number of courses	Number of lab titles
Magnetism	Force-distance relationship, Lorentz force, magnetic domains, magnetic force on conductor	11	12
Materials (simple)	Anelasticity of solids, bending a bar, deformation, elasticity, elastic torsion, elongation of a wire, plasticity of solids, Young's modulus	10	14
Mechanical oscillations	Coupled oscillators, damping, forced mechanical oscillator (with and without friction), harmonic motion, mass/spring, normal modes, resonance	12	16
Medical applications	Doppler sonography, electrocardiogram (ECG), eye optics, fluids (blood, sweat, tears), imaging sonography, myography, optical coherence tomography, optical computed tomography (CAT) scan, positron emission tomography (PET), radioactivity and health, ultrasound	14	19
Microscopy	Atomic force microscopy (AFM), evanescent light scattering, magnetic force microscopy (MFM), scanning electron microscopy (SEM), scanning probe microscopy (SPM), scanning tunneling microscopy (STM), transmission electron microscopy (TEM)	13	26
Millikan oil drop	Determining charge of electron	7	7
Nuclear magnetic resonance	Electron spin resonance, Larmor precession, nuclear magnetic resonance	10	14
Optics (advanced)	Acousto-optic modulator (AOM), Berry phase (Pancharatnam phase), critical opalescence, dynamic light scattering, Fabry-Perot cavity, fluorescence correlation spectroscopy, Fourier optics, heterodyning, ion traps, laser interference lithography, magneto-optical trap (MOT), magneto-optic effects, optical fibers (NOT fiber lasers), optical pumping, optical trapping and/or tweezers, photocarrier grating, photoluminescence quantum yield, photomultiplier tube (PMT), photon transfer functions, pump-probe (including femtosecond), quantum cryptography, quantum experiments, Raman-Nath diffraction (acousto-optic diffraction, AOD), single photon correlation, single photon detectors, sonoluminescence, spatial light modulator, wavefront shaping, Zeeman effect	17	49
Optics (intermediate)	Birefringence, diffraction, greenhouse effect, holography, interferometry and interference (including Young's experiment), microwave diffraction, microwave reflection, microwave scattering, Newton's rings, photometry, physical optics, prisms, refraction (Snell's law), schematic diagrams, spatial filtering, spectroscopy (optical), thin films, wavelengths of visible light	68	108
Optics (simple)	Alignment, building a light microscope and/or Kohler's illumination principle, geometric optics, HeNe lasers, lamps, lenses, light sources, mirrors, polarization and/or Brewster's angle, rail optics, ray optics, telescopes and/or Galileoscopes	23	39
Particle physics	accelerator physics, alpha rays, angular correlation, beta rays, chain reactions, cloud chamber, coincidence measurements, compton scattering, cosmic ray muons, cross-section, dark matter detection, Fe57 metastable state lifetime, gamma absorption and/or attenuation, gamma spectroscopy, mass of neutron, Mössbauer effect and/or spectroscopy, muon lifetime, nuclear power, particle tracking, positron emission tomography (PET), relativistic electrons, spectroscopy, strangeness, Z0 decays	18	44
Pendulum	chaotic, coupled, Kater's, physical, Pohl, reversion, simple, torsional	23	33
Photoelectric effect		9	9
Plotting	graphical presentation of measurements, graphing motion, graphing with Excel, graphs, plotting, presenting data	5	5
Radioactivity	radioactivity, half-life, attenuation	14	17
Solar cells		6	6

(Table continued)

TABLE XI. *(Continued)*

Code	Definitions	Number of courses	Number of lab titles
Spectroscopy	atomic spectra, Balmer series, dynamic light scattering, fluorescence correlation spectroscopy, Fourier transform infrared (FTIR) spectroscopy, laser spectroscopy, mass spectroscopy and/or spectrometry, optical grating spectroscopy, Raman spectroscopy, rubidium saturation spectroscopy, saturation spectroscopy, spectroscopy, time-resolved absorption spectroscopy, time-resolved fluorescence spectroscopy; NOT gamma spectroscopy, NOT Mössbauer spectroscopy	28	37
Speed of light		5	5
Speed of sound	measurement of Doppler effect, properties of sound waves, speed of sound in air and in materials	8	8
Springs	Hooke's law, spring constant	8	9
Stern-Gerlach experiment		3	3
Test and measurement equipment	calibration, drift chambers, field programmable gate arrays (FPGAs), image processing, lock-in amplifiers, microcontrollers, micrometers, oscilloscopes, Palmer caliper, periodic signals, reading seismic data, slider caliper, strain gauges, thermocouples/thermometers/thermistors, transducers, Vernier calipers	27	30
Thermodynamics	adiabatic experiments, Boltzmann constant, Boyle's law (Boyle-Mariotte law), calorimetry, critical point, evaporation in a vacuum, heat capacity, heat capacity ratio (Cp/Cv), heat conduction, heat engine, heat pump, heat transfer, heat of combustion, heat of fusion, heat of vaporization (including of liquid nitrogen), ideal gas, linear expansion, Newton's law of cooling, phase transitions, Piston effect, solar cooking box, specific heat, Stirling cycle, temperature dependence of surface tension, thermal expansion, triple points, water vapor	26	56
Viscosity	determination of viscosity of fluid, free fall of sphere in viscous fluid	9	11
Waves	electrical waves, mechanical vibrations, properties of sound waves, resonance of electromagnetic waves, sound frequency measurement, sound resonance in open-end tube, standing electromagnetic waves, standing waves, thermal waves, traveling waves, vibrating strings, vibrations, water waves	17	22
X-ray experiments	X-ray diffraction, X-ray experiments	12	15

India, Indonesia, Ireland, Israel, Kenya, Norway, Oman, Pakistan, Poland, South Africa, South Korea, and the United States. Each country had one instructor interview except for the United States, which had two (the first was to check the interview protocol as well as collect data; United States is varied enough to warrant two interviews). Because our authors are from Germany, Italy, England, and the Netherlands, we specifically did not contact instructors in these countries for interviews because our authors could provide the necessary information.

Table IX shows the number of responses to the survey we received for each country.

Next, we present in Table X the number of hours per week beyond the scheduled time that instructors estimate students spend in the lab.

Finally, we provide the definitions of all codes used to qualitatively code the lab titles in Table XI. Codes were developed emergently, where each lab title was first categorized on fine-grained scale and then codes were collapsed to create the final categories. In some cases, titles might be double or triple coded as they fall into two or three clear categories. For example, positron emission tomography (PET) is both a particle physics experiment but also one with medical applications and therefore is coded in both categories.

[1] R. Monastersky and R. Van Noorden, 150 years of Nature: A data graphic charts our evolution, Nature (London) **575**, 22 (2019).

[2] M. D. Caballero, D. R. Dounas-Frazer, H. J. Lewandowski, and M. R. Stetzer, Labs are Necessary, and We Need to Invest in Them, APS news **27** (2018), https://www.aps.org/archives/publications/apsnews/201805/backpage.cfm.

[3] Focused collection of Physical Review PER: Instructional labs-improving traditions and new directions (2022), https://journals.aps.org/prper/collections/PER-LAB.

[4] A. Hofstein and V. N. Lunetta, The laboratory in science education: Foundations for the twenty-first century, Sci. Educ. **88**, 28 (2004).

[5] J. M. May, Historical analysis of innovation and research in physics instructional laboratories: Recurring themes and future directions, Phys. Rev. Phys. Educ. Res. **19**, 020168 (2023).

[6] J. Kozminski, H. Lewandowski, N. Beverly, S. Lindaas, D. Deardorff, A. Reagan, R. Dietz, R. Tagg, M. EblenZayas, and J. Williams, AAPT recommendations for the undergraduate physics laboratory curriculum, American Association of Physics Teachers **29** (2014), https://www.aapt.org/Resources/upload/LabGuidlines-Document_EBendorsed_nov10.pdf.

[7] B. R. Wilcox, B. M. Zwickl, R. D. Hobbs, J. M. Aiken, N. M. Welch, and H. Lewandowski, Alternative model for administration and analysis of research-based assessments, Phys. Rev. Phys. Educ. Res. **12**, 010139 (2016).

[8] B. Van Dusen, M. Shultz, J. M. Nissen, B. R. Wilcox, N. G. Holmes, M. Jariwala, E. W. Close, H. J. Lewandowski, and S. Pollock, Online administration of research-based assessments, Am. J. Phys. **89**, 7 (2021).

[9] A. Madsen, S. B. McKagan, and E. C. Sayre, Resource letter RBAI-1: Research-based assessment instruments in physics and astronomy, Am. J. Phys. **85**, 245 (2017).

[10] S. G. Sireci, On validity theory and test validation, Educ. Res. **36**, 477 (2007).

[11] S. Messick, Standards of validity and the validity of standards in performance assessment, Educ. Meas. **14**, 5 (1995).

[12] L. J. Cronbach, Five perspectives on the validity argument, in *Test Validity*, edited by H. Wainer and H. I. Braun (Routledge, New York, 1988) Chap. 1, pp. 3–18.

[13] M. Kane, T. Crooks, and A. Cohen, Validating measures of performance, Educ. Meas. **18**, 5 (1999).

[14] J. K. Hemphill and C. M. Westie, The measurement of group dimensions, J. Psychol. **29**, 325 (1950).

[15] R. J. Rovinelli and R. K. Hambleton, On the use of content specialists in the assessment of criterion-referenced test item validity, Tijdschrift voor onderwijsresearch **2**, 49 (1977), https://psycnet.apa.org/record/1979-12368-001.

[16] J. C. Nunnally and I. H. Bernstein, *Psychometric Theory*, 3rd ed. (McGraw-Hill, New York, 1994).

[17] American Educational Research Association, ed., *Report and Recommendations for the Reauthorization of the Institute of Education Sciences* (American Educational Research Association, Washington, D.C., 2011).

[18] B. S. Everitt, S. Landau, M. Leese, and D. Stahl, *Cluster Analysis*, 5th ed., Wiley Series in Probability and Statistics (Wiley, Chichester, West Sussex, U.K, 2011), https://www.wiley.com/en-us/Cluster+Analysis%2C+5th+Edition-p-9780470749913.

[19] E. Etkina, When learning physics mirrors doing physics, Phys. Today **76** (10), 26 (2023).

[20] S. Chen, H.-C. Lo, J.-W. Lin, J.-C. Liang, H.-Y. Chang, F.-K. Hwang, G.-L. Chiou, Y.-T. Wu, S. W.-Y. Lee, H.-K. Wu, C.-Y. Wang, and C.-C. Tsai, Development and implications of technology in reform-based physics laboratories, Phys. Rev. ST Phys. Educ. Res. **8**, 020113 (2012).

[21] I. Kontro, O. Heino, I. Hendolin, and S. Galambosi, Modernisation of the intermediate physics laboratory, Eur. J. Phys. **39**, 025702 (2018).

[22] L. Ketonen, A. Lehtinen, and P. Koskinen, Assessment designs of instructional labs: A literature review and a design model, Phys. Rev. Phys. Educ. Res. **19**, 020601 (2023).

[23] S. Narayanan, P. Sarin, N. Pawar, and S. Murthy, Teaching research skills for experimental physics in an undergraduate electronics lab, Phys. Rev. Phys. Educ. Res. **19**, 020103 (2023).

[24] M. Alemani, The redesign of an introductory physics laboratory course, Il Nuovo Cimento C **46**, 1 (2023).

[25] G. Organtini and E. Tufino, Effectiveness of a laboratory course with Arduino and Smartphones, Educ. Sci. **12**, 898 (2022).

[26] P. Logman and J. Kautz, From dublin descriptors to implementation in bachelor labs, J. Phys. Conf. Ser. **1929**, 012065 (2021).

[27] F. Pols, One setup for many experiments: enabling versatile student-led investigations, Phys. Educ. **59**, 015007 (2023).

[28] P. Logman, Engaging theoretically primed students in second year lab courses, J. Phys. Conf. Ser. **2727**, 012021 (2024).

[29] F. R. Bradbury and F. Pols, A pandemic-resilient open-inquiry physical science lab course which leverages the maker movement, Electron. J. Res. Sci. Math. Educ. **24**, 60 (2020), https://files.eric.ed.gov/fulltext/EJ1285251.pdf.

[30] F. Bouquet, J. Bobroff, M. Fuchs-Gallezot, and L. Maurines, Project-based physics labs using low-cost open-source hardware, Am. J. Phys. **85**, 216 (2017).

[31] L. Feenstra, C. Julia, and P. Logman, A Lego ® Mach–Zehnder interferometer with an Arduino detector, Phys. Educ. **56**, 023004 (2021).

[32] P. Pirinen, A. Lehtinen, and N. G. Holmes, Impact of traditional physics lab instruction on students' critical thinking skills in a Finnish context, Eur. J. Phys. **44**, 035702 (2023).

[33] S. Z. Lahme, P. Klein, A. Lehtinen, A. Müller, P. Pirinen, L. Rončević, and A. Sušac, Evaluating digital experimental tasks for physics laboratory courses, PhyDid B **1** (2023), https://ojs.dpg-physik.de/index.php/phydid-b/article/view/1391.

[34] E. Teichmann, H. Lewandowski, and M. Alemani, Investigating students' views of experimental physics in German laboratory classes, Phys. Rev. Phys. Educ. Res. **18**, 010135 (2022).

[35] I. Bearden, L. Dvořák, and G. Planinšič, Work group 2 position paper: Experiments and laboratory work in teacher education, J. Phys. Conf. Ser. **2297**, 012008 (2022).

[36] C. F. J. Pols, H. Lewandowski, P. Logman, and F. Bradbury, Differences and similarities in approaches to physics LAB-courses, in *World Conference on Physics Education* (Hanoi, Viet Nam, 2021).
[37] N. Holmes and H. Lewandowski, Investigating the landscape of physics laboratory instruction across North America, Phys. Rev. Phys. Educ. Res. **16**, 020162 (2020).
[38] S. Salehi, C. J. Ballen, K. B. Laksov, K. Ismayilova, P. Poronnik, P. M. Ross, V. Tzioumis, and C. Wieman, Global perspectives of the impact of the COVID-19 pandemic on learning science in higher education, PLoS One **18**, e0294821 (2023).
[39] B. M. Zwickl, N. Finkelstein, and H. J. Lewandowski, Development and validation of the Colorado learning attitudes about science survey for experimental physics, AIP Conf. Proc. **1513**, 442 (2013).
[40] B. M. Zwickl, T. Hirokawa, N. Finkelstein, and H. J. Lewandowski, Epistemology and expectations survey about experimental physics: Development and initial results, Phys. Rev. ST Phys. Educ. Res. **10**, 010120 (2014).
[41] B. 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 (2016).
[42] B. R. Wilcox and H. J. Lewandowski, Students' views about the nature of experimental physics, Phys. Rev. Phys. Educ. Res. **13**, 020110 (2017).
[43] J. Henriksson, En analys av hur en undervisning med Investigative Science Learning Environment (ISLE) bör påverka elevers syn på fysik, fysikinlärning och fysikexperiment. Samt en svensk översättning av två Research-Based Assessment Instruments (RBAIs)—CLASS och ECLASS, Bachelor's thesis, Uppsala University, 2020.
[44] S. Levy, Z. Kapach, E. Magen, and E. Yerushalmi, Redefining lab norms via professional learning communities of physics teachers, presented at PER Conf. 2020, virtual conference, 10.1119/perc.2020.pr.Levy.
[45] C. Walsh, K. N. Quinn, C. Wieman, and N. Holmes, Quantifying critical thinking: Development and validation of the physics lab inventory of critical thinking, Phys. Rev. Phys. Educ. Res. **15**, 010135 (2019).
[46] C. Walsh, K. N. Quinn, and N. G. Holmes, Assessment of critical thinking in physics labs: Concurrent validity, presented at PER Conf. 2018, Washington, DC, 10.1119/perc.2018.pr.Walsh.
[47] K. N. Quinn, C. E. Wieman, and N. G. Holmes, Interview validation of the Physics Lab Inventory of Critical thinking (PLIC), presented at PER Conf. 2018, Cincinnati, OH, 10.1119/perc.2017.pr.076.
[48] N. G. Holmes and C. E. Wieman, Assessing modeling in the lab: Uncertainty and measurement, in *Proceedings of the 2015 Conference on Laboratory Instruction Beyond the First Year* (American Association of Physics Teachers, College Park, MD, 2015), pp. 44–47.
[49] R. Teodorescu, C. Bennhold, and G. Feldman, Enhancing cognitive development through physics problem solving: A taxonomy of introductory physics problems, AIP Conf. Proc. **1064**, 203 (2008).
[50] R. E. Teodorescu, C. Bennhold, G. Feldman, and L. Medsker, New approach to analyzing physics problems: A taxonomy of introductory physics problems, Phys. Rev. ST Phys. Educ. Res. **9**, 010103 (2013).
[51] B. R. Wilcox and H. Lewandowski, Open-ended versus guided laboratory activities:Impact on students' beliefs about experimental physics, Phys. Rev. Phys. Educ. Res. **12**, 020132 (2016).
[52] V. Borish, A. Werth, N. Sulaiman, M. F. Fox, J. R. Hoehn, and H. Lewandowski, Undergraduate student experiences in remote lab courses during the COVID-19 pandemic, Phys. Rev. Phys. Educ. Res. **18**, 020105 (2022).
[53] M. F. Fox, J. R. Hoehn, A. Werth, and H. Lewandowski, Lab instruction during the COVID-19 pandemic: Effects on student views about experimental physics in comparison with previous years, Phys. Rev. Phys. Educ. Res. **17**, 010148 (2021).
[54] A. Werth, J. R. Hoehn, K. Oliver, M. F. Fox, and H. Lewandowski, Instructor perspectives on the emergency transition to remote instruction of physics labs, Phys. Rev. Phys. Educ. Res. **18**, 020129 (2022).
[55] J. van Den Akker, W. Kuiper, and U. Hameyer, *Curriculum Landscapes and Trends* (Springer Netherlands, Dordrecht, 2003).
[56] A. C. McCormick and V. M. H. Borden, Higher education institutions, types and classifications of, in *The International Encyclopedia of Higher Education Systems and Institutions*, edited by P. N. Teixeira and J. C. Shin (Springer Netherlands, Dordrecht, 2020), pp. 697–705.
[57] A. Usher and J. Medow, *Global Higher Education Rankings 2010: Affordability and Accessibility in Comparative Perspective* (Higher Education Strategy Associates, Stafford, Virginia, 2010).
[58] P. G. Altbach, L. Reisberg, and L. E. Rumbley, *Trends in Global Higher Education: Tracking an Academic Revolution* (Brill, Leiden, The Netherlands, 2010).
[59] F. Ziegele, Classification of Higher Education Institutions: The European Case, Pensam. Educ. **50**, 76 (2013), https://openurl.ebsco.com/EPDB%3Agcd%3A1%3A5119329/detailv2?sid=ebsco%3Aplink%3Ascholar&id=ebsco%3Agcd%3A90604622&crl=f.
[60] A. C. McCormick and C.-M. Zhao, Rethinking and reframing the carnegie classification, Change **37**, 51 (2005).
[61] L. A. Goodman, Snowball sampling, Ann. Math. Stat. **32**, 148 (1961).
[62] P. Biernacki and D. Waldorf, Snowball sampling: Problems and techniques of chain referral sampling, Sociol. Methods Res. **10**, 141 (1981).
[63] L. Dana, B. Pollard, and S. Mueller, Professional contexts of physics instructional labs: More than technical support, Phys. Rev. Phys. Educ. Res. **19**, 020127 (2023).
[64] Advanced laboratory physics association (ALPhA), https://advlab.org [accessed March 25, 2024].
[65] American physical society forum on education (), https://engage.aps.org/fed/home [accessed March 25, 2024].
[66] American physical society topical group on physics education research (), https://engage.aps.org/gper/home [accessed March 25, 2024].
[67] Jiscmail—physics education, https://www.jiscmail.ac.uk/cgi-bin/webadmin?A0=PHYSICS-EDUCATION [accessed March 25, 2024].

[68] Lab taxonomy survey, qualtrics, https://cuboulder.qualtrics.com/jfe/form/SV_ehWi7ucbFNk1kFM [accessed January 10, 2024].
[69] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevPhysEducRes.20.020117 for an adaptation of the Qualtrics version of the full lab taxonomy survey.
[70] Guide to instructional laboratories and experimental skills, https://ep3guide.org/guide-overview/instructional-laboratories-and-experimental-skills [accessed January 18, 2023].
[71] B. M. Zwickl, N. Finkelstein, and H. Lewandowski, The process of transforming an advanced lab course: Goals, curriculum, and assessments, Am. J. Phys. **81**, 63 (2013).
[72] E. Etkina, A. Karelina, and M. R. Villasenor, Studying transfer of scientific reasoning abilities, AIP Conf. Proc. **883**, 81 (2007).
[73] R. J. Beichner, J. M. Saul, D. S. Abbot, J. J. Morse, D. L. Deardorff, R. J. Allain, S. W. Bonham, M. H. Dancy, and J. S. Risley, The student-centered activities for large enrollment undergraduate programs (scale-up) project, in *Reviews in PER Volume 1: Research-Based Reform of University Physics*, edited by E. Redish and P. Cooney (American Association of Physics Teachers, College Park, MD, 2007).
[74] E. Brewe, Modeling theory applied: Modeling Instruction in introductory physics, Am. J. Phys. **76**, 1155 (2008).
[75] B. Pollard, R. Hobbs, R. Henderson, M. D. Caballero, and H. Lewandowski, Introductory physics lab instructors' perspectives on measurement uncertainty, Phys. Rev. Phys. Educ. Res. **17**, 010133 (2021).
[76] M. Vignal, G. Geschwind, B. Pollard, R. Henderson, M. D. Caballero, and H. Lewandowski, Survey of physics reasoning on uncertainty concepts in experiments: An assessment of measurement uncertainty for introductory physics labs, Phys. Rev. Phys. Educ. Res. **19**, 020139 (2023).
[77] M. F. J. Fox, B. Pollard, L. Ríos, and H. J. Lewandowski, Capturing modeling pathways using the Modeling Assessment for Physics Laboratory Experiments, presented at PER Conf. 2020, virtual conference, 10.1119/perc.2020.pr.Fox.
[78] Google Translate, https://translate.google.com/.
[79] J. Kant, Simple word cloud generator, available at https://www.simplewordcloud.com.
[80] A. Werth, K. Oliver, C. G. West, and H. Lewandowski, Assessing student engagement with teamwork in an online, large-enrollment course-based undergraduate research experience in physics, Phys. Rev. Phys. Educ. Res. **18**, 020128 (2022).
[81] M. Sundstrom, D. G. Wu, C. Walsh, A. B. Heim, and N. Holmes, Examining the effects of lab instruction and gender composition on intergroup interaction networks in introductory physics labs, Phys. Rev. Phys. Educ. Res. **18**, 010102 (2022).
[82] E. M. Stump, M. Dew, S. Jeon, and N. Holmes, Taking on a manager role can support women's physics lab identity development, Phys. Rev. Phys. Educ. Res. **19**, 010107 (2023).
[83] M. Dew, E. Hunt, V. Perera, J. Perry, G. Ponti, and A. Loveridge, Group dynamics in inquiry-based labs: Gender inequities and the efficacy of partner agreements, Phys. Rev. Phys. Educ. Res. **20**, 010121 (2024).
[84] C. F. J. Pols, P. J. J. M. Dekkers, and M. J. de Vries, Integrating argumentation in physics inquiry: A design and evaluation study, Phys. Rev. Phys. Educ. Res. **19**, 020170 (2023).
[85] L. Ding and R. Beichner, Approaches to data analysis of multiple-choice questions, Phys. Rev. ST Phys. Educ. Res. **5**, 020103 (2009).
[86] Arbeitsgruppe physikalische praktika (AGPP), https://www.dpg-physik.de/vereinigungen/fachuebergreifend/ag/agpp [accessed April 18, 2024].
[87] M. F. J. Fox, B. M. Zwickl, and H. J. Lewandowski, Preparing for the quantum revolution: What is the role of higher education?, Phys. Rev. Phys. Educ. Res. **16**, 020131 (2020).
[88] Groupe international de recherche sur l'enseignement de la physique—girep (), https://www.girep.org [accessed March 25, 2024].
[89] Girep thematic group on laboratory based teaching in physics (LabTiP) (), https://www.girep.org/thematic-groups/laboratory-based-teaching-in-physics/ [accessed April 19, 2024].