PHYSICAL REVIEW PHYSICS EDUCATION RESEARCH **19,** 020166 (2023)

# Improving scientific abilities through lab report revision in a high school investigative science learning environment classroom

Danielle Buggé

*Graduate School of Education, Rutgers University,*
*10 Seminary Place, New Brunswick, New Jersey 08901, USA*

(Received 29 May 2023; accepted 27 November 2023; published 15 December 2023)

[This paper is part of the Focused Collection on Instructional labs: Improving traditions and new directions.] National organizations set goals of engaging students in experimentation and authentic scientific reasoning while developing normative concepts to help them develop essential skills and competencies necessary to succeed in our rapidly changing world. One framework that meets these goals is the investigative science learning environment (ISLE) approach to teaching and learning physics. While studies have been conducted on student development of scientific abilities in ISLE-based classrooms, little is known about how formative assessment helps promote student growth in this setting. This paper presents on the findings from an empirical study on how the revision of written laboratory reports positively impacts first-year high school student development of scientific abilities in an ISLE-approach classroom. We argue that the opportunity to revise written work provides students with additional exposures that are necessary to support their scientific ability development. Furthermore, we explore positive correlations between student involvement in the collaborative writing process and attitudes regarding experimental physics.

DOI: 10.1103/PhysRevPhysEducRes.19.020166

## I. BACKGROUND

Today's students are preparing for life and a career in a constantly changing world. Learning how to follow processes and procedures no longer adequately trains individuals to become productive, contributing members in the 21st century. Society needs scientifically literate citizens that understand how to think critically. Students need to develop abilities and skills that are applicable across many fields [1–5]. They need to be observant. To pose questions, develop and test explanations, make judgments about the outcomes of those experiments, and solve complex problems that do not have one right answer. They are expected to take ownership of their learning and collaborate and work effectively with their peers.

With reform efforts to science education [6–9], physics courses are a natural place to engage in such practices and develop abilities essential for success in the 21st century. Most research surrounding student performance in the laboratory has been conducted at the college level [10] and focused on the areas of scientific ability development [11,12], systematic error and critical thinking skills [13,14], understanding of measurement and uncertainty [15], feedback [16], and attitudes [17,18]. In the K–12 setting, researchers have studied student engagement with authentic science practices [19–23].

As a result, experimental activities that place greater emphasis on reasoning and practice development have an increased presence in classrooms both at the high school and college levels [10,24–27]. These experiences should help students develop habits of mind, strategies, enthusiasm, and confidence [6,18,28–30].

One approach to learning and instruction that engages students in the appropriate practices is the theoretical framework and curriculum materials that make up the investigative science learning environment approach, or the ISLE approach [31,32]. This student-centered approach to learning and teaching meets the recommendations of organizations at the K–12 and college levels [5–9]. ISLE has two core intentionalities: students learning physics by engaging in reasoning processes similar to those of physicists, and that their well being is enhanced by the experience [33]. Enhancing student well being means teaching them that intelligence is not fixed [34,35] and persistence and motivation [36,37] are equally important to learning physics. Directed feedback [38–40] and multiple formative assessment opportunities [41–44] will help students think like physicists and build habits of minds for 21st century careers [1,2,4].

### A. Overview of the ISLE approach

In an ISLE-based classroom, students engage in the development of physics ideas through reasoning processes

---

TABLE I. Sample ISLE investigations conducted during the mechanics portion of the school year.

| Type of experiment | Example |
|---|---|
| Observational experiment | Water design lab<br>Observe and find a pattern: Equipment: bucket of water, bottled filled with sand, spring scale, ruler. Design an experiment that will allow you to answer some questions about forces. Focus on the magnitudes and directions of forces that the water will exert on the bottle… |
| Testing experiment | Fan cart lab<br>Design an experiment to test whether the fan cart moves at constant speed or constant acceleration. Make measurable prediction(s) for the outcome of your experiment(s) based on these two models of motion.<br>Perform the experiment, record the results, and decide which model you could not disprove and which model you were able to disprove. |
| Application experiment | Mu shoe lab<br>Determine the maximum coefficient of static friction between your shoe and a surface using *two independent methods*. Only *one* method may use the spring scale. |

similar to those physicists use to construct and apply knowledge [32]. Students consistently collaborate with their peers to develop, test, and use physical models in a variety of contexts. Every new concept starts with a series of observational experiments. Students look for patterns and come up with possible explanations (or hypotheses). Next, students need to determine the plausibility of these hypotheses. To do this, the students temporarily accept the hypotheses as true and design experiments whose outcome they predict using the hypotheses under test. After performing the testing experiments, students make a judgment about the hypothesis. If the experimental and predicted outcomes do not align, students need to revise their assumptions or reject the hypothesis. If the experimental and predicted outcomes agree, the students may further test the hypothesis or apply this knowledge to answer new questions and solve practical problems [31]. An example of each type of experiment is provided in Table I. This process of knowledge development and application through experimentation is integrated into all aspects of the learning (not just the laboratory) and incorporates natural opportunities for students to develop scientific abilities.

### B. Scientific abilities and rubrics

Scientific abilities are the processes, procedures, and methods used by physicists as they construct and apply knowledge. They were constructed based on the history of practices in physics [45], goals of the physics curriculum [46], recommendations of cognitive scientists [47], and analysis of science process test items [48]. It is necessary for students to develop these abilities in order to productively contribute to a scientific community. In reformed courses, these practices are frequently included as general goals in the laboratory and evaluated using scientific ability rubrics described below. These abilities are as follows:

1. the ability to represent information in multiple ways,
2. the ability to design and conduct an experiment to investigate a phenomenon,
3. the ability to design and conduct a testing experiment (testing an idea, hypothesis or explanation, or mathematical relation),
4. the ability to design and conduct an application experiment,
5. the ability to communicate scientific ideas,
6. the ability to collect and analyze experimental data, and
7. the ability to evaluate models, equations, solutions, and claims [46].

It takes time for students to develop scientific abilities using a learning framework such as ISLE [12]. Students need constructive feedback, opportunities for improvement, and repeated exposure to each of the abilities in order to develop a capacity for metacognitive thinking [41,43,49–52].

Within ISLE approach classrooms, students frequently engage in writing tasks. One of the best places to do science writing in physics classes is in the laboratory [53]. Historically, writing to learn was tied to the development of laboratory methods in college science classes [54]. With recent reform efforts, instructors are moving away from using lab reports to show completion, and instead utilizing scientific writing as a way to drive instruction in the classroom [55]. Writing provides students the opportunity to not only share out well-formulated ideas but also a space to construct knowledge, synthesize ideas, and reflect on their learning and understanding [56–58]. Furthermore, the act of writing makes student thinking visible [57] for their peers which is essential in a collaborative learning environment. Within the K–12 [59,60] and college [61–63] classrooms, most scientific writing takes place in lab classes [64].

ISLE approach investigations are formative [41] and the framework encourages students to revise their work. Revision is an essential part of the writing process [65] and shifts the emphasis from a grade toward understanding [41].

TABLE II. Sample scientific subability rubric for identifying assumptions.

| Scientific ability | Missing | Inadequate | Needs improvement | Adequate |
|---|---|---|---|---|
| Is able to identify assumptions made in using the mathematical procedure. | No attempt is made to identify assumptions. | An attempt is made to identify assumptions, but the assumptions are irrelevant or incorrect for the situation. | Relevant assumptions are correctly identified but are not significant for solving the problem. | All relevant assumptions are correctly identified. |

While there are several studies on the use of formative assessment in physics classrooms [65–69], research on the effects of revisions on student growth and understanding in any classroom is limited [19,70,71]. ISLE activities are designed so that students have the opportunity to take instructor feedback and make changes to their written work without penalty, similar to the peer review process in the production of scientific knowledge.

To assist students with scientific ability development, each of the seven abilities is broken down into smaller subabilities that students focus on when conducting an investigation. Rubrics are used to lay out the expectations for each scientific subability. The scientific abilities rubrics are available online (see Ref. [72]). Table II contains a sample rubric for the subability to identify assumptions made in using the mathematical procedure. Assumptions are factors that are assumed to be true when applying a certain mathematical model to a real-world situation [28].

All scientific ability rubrics contain evaluation criteria for assessing the goals, descriptors for each level of achievement, and a consistent scoring strategy [73–75]. The detailed explanations in the rubric item subability descriptors are used to identify current student progress toward proficiency [76,77]. The scientific ability rubrics make the learning goals for each task visible so students can use these to self-assess their progress. Additionally, instructor feedback is provided by identifying the current level of student performance, which is either adequate (student performance is at the expected level), needs improvement (student performance is on the right track but needs some work), inadequate (student performance attempts to meet the standards for the ability but does not), and missing (students do not attempt to demonstrate understanding of the ability).

TABLE III. Sample student responses for “is able to identify assumptions made in using the mathematical procedure”.

| Rubric score | Sample student response | Discussion |
|---|---|---|
| Missing | This is all under the assumption that the fan cart would be unaffected by other uncertainties, such as gusts of wind. | The student does not make any assumptions that are relevant for solving the problem. |
| Inadequate | We assume that the cart is moving in a straight line on flat ground. | The student attempts to make assumptions. |
| Needs improvement | It is assumed that the cart is going perfectly straight the entire time.<br>It is assumed that the floor is perfectly flat.<br>It is assumed that the direction flow of air in the room is negligible. | The student makes relevant assumptions that are important, but not significant for solving the problem as it relates to the mathematical procedure. |
| Adequate | We made some assumptions during this lab while conducting the experiment and analyzing the data. We assumed that the car is going perfectly straight and that the only thing causing its motion was the fan. While calculating the acceleration with the manual data, I assumed the car had constant acceleration. Without constant acceleration, the instantaneous velocity at the midpoint of the two times cannot be approximated by the average velocity between the two times. | The student makes relevant assumptions for solving the problem. This adequate response makes mention of other interactions that could affect the motion of the fan as well as that the cart moves with constant acceleration (after they’ve determined this is the motion) and the implication of this assumption on the mathematical model. |

Sample student responses for the subability “is able to identify assumptions made in using the mathematical procedure” are shown in Table III.

Typically, investigations focus on four to six subabilities and students are evaluated on the same rubric items at multiple points in the school year. Students have the opportunity to revise their work. At the college level, it takes approximately eight weeks for students in introductory physics courses to begin to reach 80% proficiency with many of the scientific subabilities [12].

At the high school level, it takes students a longer period of time to develop these abilities, but their performance often surpasses that of introductory college students [20].

However, educators are not aware of *how* the opportunity to revise written laboratory reports helps high school students in a first-year introductory physics course develop the ability to think like a scientist. We are unaware of how individual students contribute when writing up findings collaboratively. Furthermore, we are unaware of the effect of scientific ability development on high school students’ attitudes regarding experimental physics.

The study reported in this paper addresses the following research questions:

1. How does the opportunity to revise written laboratory reports affect whole class development of scientific abilities?
2. What is the effect of scientific ability development in an ISLE-approach classroom on student attitudes toward and confidence with experimental physics?

## II. METHODOLOGY

We used a convergent parallel mixed-methods approach [78] to explore the effects of the opportunity to revise on high school student development of science-process abilities and attitudes regarding experimental physics in a first-year physics course.

### A. Setting

This study took place in a Northeast suburban public high school over a consecutive three-year period. At the time of the study, there were four physics instructors at this school who were trained in and implemented the ISLE approach in their classrooms; two of these instructors’ classrooms were sampled. The student population in this high school was 70% Asian, 20% White, 5% Black, and 5% Hispanic. All students were provided a Chromebook and school Google account. There was a high expectation for excellence and achievement in the district from parents and students.

This study was conducted in ten sections of a first-year honors physics course (15- to 18-year-old students enrolled in the course) that followed the ISLE approach. The two course instructors utilized ISLE-based instructional materials including physics union mathematics [79] curricular resources, scientific abilities rubrics [72], and the active learning guide [80].

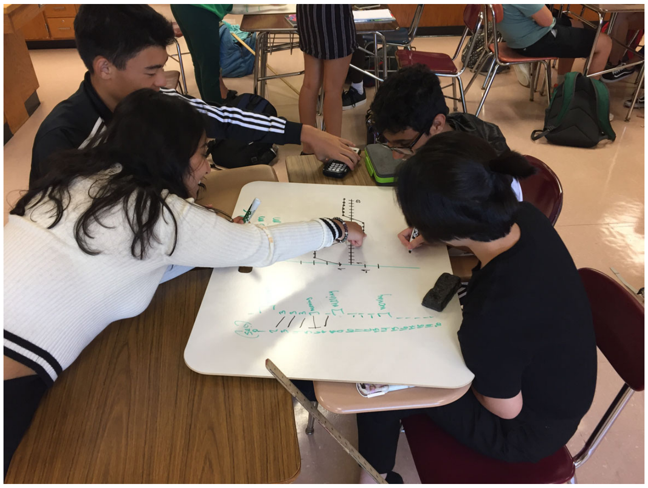


FIG. 1. Students working collaboratively in the classroom.

Enrollment was on average 24 students per class and class periods were either 60 or 80 min depending on the day in the rotation. Total instructional time ranged from 200 to 280 min per week. The physical layout of the classroom lent itself to collaboration as students are seated in groups. These groups changed approximately once a month so students had an opportunity to learn with all of their peers over the course of the school year. Figure 1 shows students working together in class on an investigation. Within these groups, students engaged with the NGSS science practices [8] and scientific abilities [46] as they worked together to develop, test, and apply physics concepts to the real world.

### B. Sampling

The students were sampled purposefully [81] from all honors physics sections offered at the high school each year (Table IV). Over a three-year period, we collected work from 230 different students in ten sections of honors physics. Fifty-five percent of these students identified as male while 45% identified as female. During year 1 and year 2 work was collected from a single classroom. During year 3, work was collected from two classrooms. Instructor 1 identified as female and instructor 2 identified as male. Both instructors were graduates of the same ISLE-based teacher preparation

TABLE IV. Overview of sections and students sampled for the study.

| Year | Number of sections offered | Number of sections sampled | Number of students | Number of classroom instructors |
|---|---|---|---|---|
| Year 1 | 8 | 3 | 69 | 1 |
| Year 2 | 7 | 2 | 46 | 1 |
| Year 3 | 7 | 5 | 115 | 2 |

TABLE V. Average FCI Scores for sampled sections.

| Section | Pre | Post | Gain |
|---|---|---|---|
| Year 1 | $9.10 \pm 4.38$ | $19.69 \pm 5.43$ | $0.53 \pm 0.28$ |
| Year 2 | $10.22 \pm 6.13$ | $19.61 \pm 5.66$ | $0.50 \pm 0.43$ |
| Year 3, instructor 1 | $9.17 \pm 4.51$ | $19.02 \pm 5.68$ | $0.50 \pm 0.30$ |
| Year 3, instructor 2 | $10.17 \pm 4.18$ | $20.02 \pm 5.35$ | $0.52 \pm 0.20$ |
| Average | $9.65 \pm 4.38$ | $19.59 \pm 5.30$ | $0.51 \pm 0.30$ |

program and had taught the honors physics course for over six years.

The Force Concept Inventory [48] was administered at the beginning and end of the mechanics unit to learn how similar or different the sampled students were for instructor 1 at the start of each school year and instructor 2 during year 3. An average normalized gain [82] was used to determine the FCI scores for each class (Table V). Whole class pre- and post-test scores are within one point of one another; the gains are also similar. These scores are on the medium to high end of reformed instruction scores [82]. This means that all four populations sampled started at the same place with their physics content knowledge.

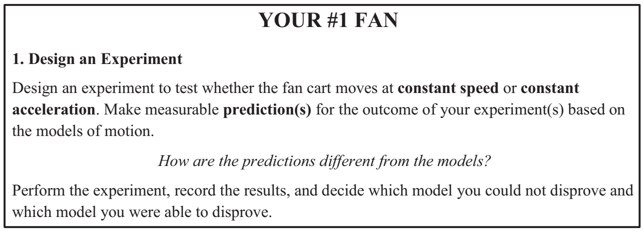
### C. Implementation of ISLE in the classroom

This study focuses on student development of science-process abilities through their written laboratory reports. In this section, we describe the process the students went through when designing an investigation, writing up their findings, self-assessing, and revising their work based on teacher feedback. Within each unit of study, the students engaged in observational, testing, and application experiments.

#### *1. Laboratory reports*

The students in this environment wrote formal laboratory reports for one or two investigations per unit of study (such as kinematics, Newton's laws, circular motion, etc.). They conducted the investigations in groups of three or four. Over the three-year period during which data were collected, the students wrote up their reports either individually or in groups depending on the year, the task, and the instructor. In year 1, instructor 1 required one class to write an individual report and two classes to write group reports for each investigation. The instructor wanted all students to have both individual and collaborative experience engaging with the scientific abilities and writing up their findings. In year 2, all lab reports except for lab 2 were written collaboratively. This lab was performed early in the year and the instructor wanted to make sure all students were comfortable writing up findings from an investigation.

In year 3, all reports for instructor 1 were written collaboratively and only lab 1 for instructor 2 was written individually. Instructor 2 also wanted to ensure that all students were comfortable writing up their findings before writing group lab reports. Recall that laboratory investigations are used to drive learning in the classroom and students design and perform these experiments collaboratively. For these classroom Instructors, writing group reports is a natural extension of the learning process. For each investigation, the students were provided an assignment sheet that included learning goals, scientific subability rubrics, and process-oriented guiding questions. Figure 2 shows the beginning of a sample investigation assignment sheet for a kinematics unit testing experiment. The full investigation is available as Supplemental Material [83].



**YOUR #1 FAN**

**1. Design an Experiment**

Design an experiment to test whether the fan cart moves at **constant speed** or **constant acceleration**. Make measurable **prediction(s)** for the outcome of your experiment(s) based on the models of motion.

*How are the predictions different from the models?*

Perform the experiment, record the results, and decide which model you could not disprove and which model you were able to disprove.

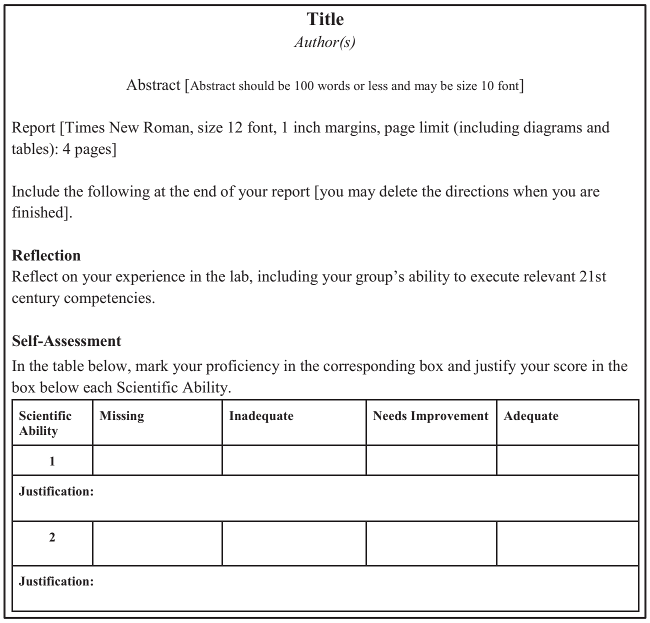
FIG. 2. Excerpt from directions for the fan cart investigation.

This modified physics union mathematics [79] investigation is an example of a testing experiment. At the end of the kinematics unit of study, students were tasked with applying their knowledge of motion to the movement of a fan cart. They had to design an experiment to test two competing models: the cart moved at constant velocity or constant acceleration. This was an appropriate testing experiment to perform at this point in the curriculum because the students had not yet learned about forces; it was reasonable for them to think the cart would move with constant motion. Students had two class periods to design the investigation, make the predictions of the outcomes based on those two different hypotheses, and collect and analyze their data. After completing their analysis, the students wrote up their findings in a report. The students were provided a teacher-created template document through Google Classroom (Fig. 3).

This template does not provide the students with guidance on how to write their report; it only contains space for their work, reflection, and self-evaluation. Providing a template ensured that the classroom teacher had access to student progress and the researcher was able to efficiently analyze the data. Figure 4 contains an excerpt from one group's report. The full report is available as Supplemental Material [83].

#### *2. Self-assessment*

After writing up their findings, each group was required to self-assess their work prior to submission. The students had to evaluate their written report using several scientific subability rubrics and to provide a commentary justifying why they believed that they deserved a particular score. As seen in the sample report (Supplemental Material [83]), the majority of students completed the self-assessment table at the end of the report. Some students chose to add comments



**Title**

*Author(s)*

Abstract [Abstract should be 100 words or less and may be size 10 font]

Report [Times New Roman, size 12 font, 1 inch margins, page limit (including diagrams and tables): 4 pages]

Include the following at the end of your report [you may delete the directions when you are finished].

**Reflection**

Reflect on your experience in the lab, including your group's ability to execute relevant 21st century competencies.

**Self-Assessment**

In the table below, mark your proficiency in the corresponding box and justify your score in the box below each Scientific Ability.

| **Scientific Ability** | **Missing** | **Inadequate** | **Needs Improvement** | **Adequate** |
|---|---|---|---|---|
| **1** | | | | |
| **Justification:** | | | | |
| **2** | | | | |
| **Justification:** | | | | |

FIG. 3. Teacher provided template for written laboratory reports. Full template is available as Supplemental Material [83].

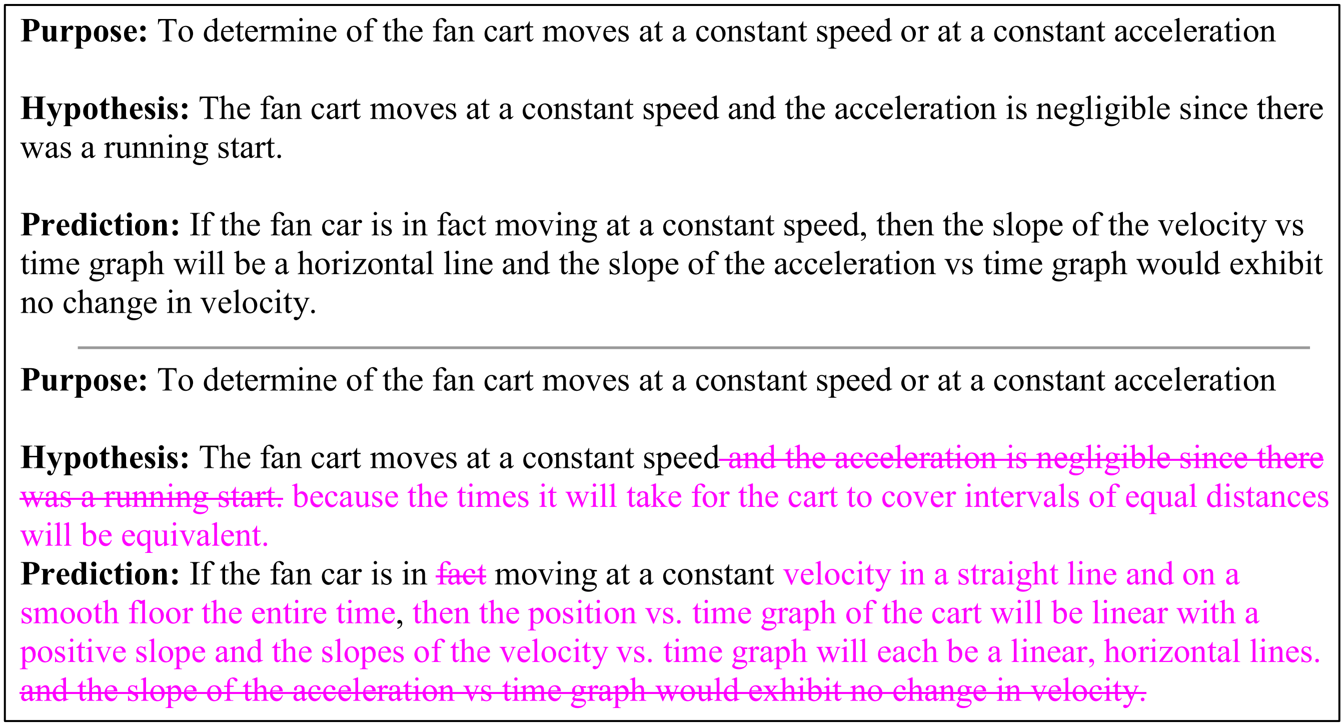
directly into the text where they felt they met the expectations of each scientific ability. Both approaches are beneficial to the classroom instructor as they provide insight into student understanding of the expectations for each scientific ability.

### *3. Teacher feedback*

Once students submitted their laboratory reports, the classroom teacher added comments and rubric scores to the assignment. The purpose of this feedback was to help students learn how to work with the scientific abilities, not provide them the correct text to include in their lab report. Teacher feedback pointed out both areas of strengths as well as areas for improvement. Feedback ranged from corrections to their scientific language, to comments about methods, to questions designed to guide student thinking. As the year progressed, and students became more comfortable working with the scientific subability rubrics, teacher feedback transitioned from more detailed in-text comments to rubric scores.

### *4. Revisions*

After receiving teacher feedback, the students had the opportunity to revise and resubmit their reports. This was optional, but the majority of students and student groups chose to revise their work. The students worked collaboratively with their peers, asked the instructor clarifying questions, and made changes to their submission in a different color. The different color text allowed for easy teacher review of their changes. A sample copy of a revised report is available as Supplemental Material [83]. The classroom teacher reevaluated the student work, added new scores to the rubric and updated the grade for the assignment.



**Purpose:** To determine of the fan cart moves at a constant speed or at a constant acceleration

**Hypothesis:** The fan cart moves at a constant speed and the acceleration is negligible since there was a running start.

**Prediction:** If the fan car is in fact moving at a constant speed, then the slope of the velocity vs time graph will be a horizontal line and the slope of the acceleration vs time graph would exhibit no change in velocity.

---

**Purpose:** To determine of the fan cart moves at a constant speed or at a constant acceleration

**Hypothesis:** The fan cart moves at a constant speed~~and the acceleration is negligible since there was a running start.~~ because the times it will take for the cart to cover intervals of equal distances will be equivalent.

**Prediction:** If the fan car is in ~~fact~~ moving at a constant velocity in a straight line and on a smooth floor the entire time, then the position vs. time graph of the cart will be linear with a positive slope and the slopes of the velocity vs. time graph will each be a linear, horizontal lines. ~~and the slope of the acceleration vs time graph would exhibit no change in velocity.~~

FIG. 4. Sample group written report for fan cart lab. Top is the initial submission; bottom is the revised submission after receiving teacher feedback. Revisions are in pink. The full report is available as Supplemental Material [83].

TABLE VI. Summary of the experiments for which students wrote laboratory reports during the mechanics portion of the school year. Level of difficulty was determined based on both cognitive and procedural demands of the task by researchers involved in development of and instruction with ISLE curricular resources.

| Lab No. | Title | Type of experiment | Summary | Level of difficulty |
|---|---|---|---|---|
| 1 | Two cars meet | Application | Predict where two cars will meet if they are initially placed 3 m apart and released at the same time. | 1–2 |
| 2 | Fan cart | Testing | Design an experiment to test whether the fan cart moves at constant speed or constant acceleration. | 4 |
| 3 | Freefall | Observational or testing | Design an experiment to investigate *how* objects move when nothing holds them up. Examine the materials available to your lab group. You may design an observational experiment or a testing experiment if you have a hypothesis. | 2 |
| 4 | Interactions with water | Observational | Observe and find a pattern: Equipment: bucket of water, bottled filled with sand, spring scale, ruler. Design an experiment that will allow you to answer some questions about forces. Focus on the magnitudes and directions of forces that the water will exert on the bottle… | 3 |
| 5 | Mu shoe | Application | Determine the maximum coefficient of static friction between your shoe and a surface using *two independent methods*. Only *one* method may use the spring scale. | 3 |
| 6 | Impulse-momentum | Testing | Design an experiment to *test* the impulse-momentum relationship. | 4–5 |
| 7 | Spring into action | Testing | You have a spring and a vertically mounted metal pole and other regular lab equipment… Design an experiment to test work-energy principle using available equipment. | 3 |
| 8 | Forensic physics | Application | Devise a procedure to determine the coefficient of friction between the tire rubber and the asphalt using a collision process. | 5 |
| 9 | Design your own experiment | Observational and testing | Design an experiment that will allow you to answer some question(s) using the physics knowledge and lab skills you've developed in the first half of the school year. Which scientific abilities are relevant for your investigation? | 5 |

### D. Data collection

Classroom artifacts include student written laboratory reports, pre- and postsurveys, and overall class performance.

#### 1. Laboratory reports

The primary data source for this exploration was student written laboratory reports from the mechanics portion of the curriculum. These artifacts were collected in order to answer research question 1. The difficulty of the investigations varied over the year with the last lab being more difficult than the first. Table VI contains a brief overview of each investigation and the level of difficulty relative to the other investigations.

The level of difficulty was determined based on both cognitive and procedural demands of the task by researchers involved in the development of and instruction through of ISLE curricular resources.

Students in instructor 1's classroom had the opportunity to revise their written reports after receiving feedback. In total, we collected 631 individual and group reports. The students revised 296 of those. Table VII highlights how many individual and group reports (initial and revised) per investigation were available for analysis. The students in instructor 2's classroom did not have the opportunity to revise their reports.

#### 2. Pre-post surveys

We also collected data from surveys that the students took at the beginning and end of each school year in order to answer research question 2. These surveys include the Colorado Learning Attitudes about Science Survey for Experimental Physics (E-CLASS) [84] and a short survey in the semantic differential response format on student perceived confidence with the scientific abilities and elements of the ISLE approach. The survey was created

TABLE VII. Laboratory investigation artifacts by year. Note that bold numbers indicate the number of reports written by groups. Regular text numbers indicate the number of reports written individually. Italic numbers (in parentheses) indicate the number of revised reports. Each student or group of students in the classroom of instructor 1 had one opportunity to revise their work for each investigation. The total number of reports scored by year are included at the bottom of the table. During year 3, lab 3 was conducted in lieu of lab 2; after year 1, labs 7 and 8 were combined into a single investigation; lab 9 was not introduced until year 2.

| Investigation | Year 1, instructor 1 | Year 2, instructor 1 | Year 3, instructor 1 | Year 3, instructor 2 |
|---|---|---|---|---|
| Lab 1 | **12** (***12***)<br>22 (*21*) | **13** (***12***) | **12** (***12***) | 66 |
| Lab 2 | **12** (***9***)<br>24 (*14*) | 44 (*18*) | | |
| Lab 3 | | | **12** (***12***) | **18** |
| Lab 4 | **12** (***8***)<br>22 (*17*) | **13** (***6***)<br>1 | **12** | **18** |
| Lab 5 | **14** (***6***)<br>22 (*13*) | **13** (***8***) | **13** (***9***) | **17** |
| Lab 6 | **10** (***10***)<br>23 (*19*) | **10** (***9***) | **13** (***13***) | **15** |
| Lab 7 | **12** (***9***)<br>23 (*7*) | | | |
| Lab 8 | **19** (***10***) | **12** (***7***) | **12** (***11***) | **17** |
| Lab 9 | | 46 (*16*) | **13** (***8***) | **11** |
| Total | 227 (*155*) | 152 (*76*) | 87 (*65*) | 165 |

by the classroom teacher (see the Appendix, Table XIII). It asked students to rank their current level of comfort and confidence with the scientific abilities such as differentiating between a hypothesis and a prediction, taking assumptions into account, making a judgment about the outcome of an experiment, and performing the three different types of experiments (observational, testing, application). Instructor 1 distributed this survey for three years prior to the onset of data collection and during the three years of the study.

### E. Data analysis

We used a microgenetic approach to analyze the laboratory reports collected from the high school classroom [85]. A microgenetic approach focuses on "the moment-by-moment change examined within a relatively short span of time for an increased number of meetings" [86] (p. 3). This approach provides researchers the opportunity to examine student work at multiple points in the year, not only the end product. We scored all reports on initial submission and after revisions. This analysis method helped us answer the research question 1 of how the opportunity to revise written laboratory reports affects whole class development of scientific abilities. Furthermore, we used a case study analysis to explore how individual students contributed to written laboratory reports. We also measured gains on the pre- and postsurveys in order to explore relationships between scientific ability development through revisions and student attitudes in the classroom. Finally, we ran a comparative analysis to explore the association between whole-class growth and factors such as whole-class attitudes, and frequency of contributions to the lab report and revisions.

#### *1. Laboratory reports*

The primary coding scheme for the lab report artifacts follows the descriptors of the subabilities on the scientific ability rubrics. The categorical values from the rubrics were converted into numerical responses (0, missing; 1, inadequate; 2, needs some improvement; and 3, adequate). The lab reports were rescored by researchers familiar with ISLE approach laboratory investigations. This was necessary because although the classroom instructors were trained in how to use the rubrics, they did not establish the reliability procedures required for research when assigning grades during the instructional process. Rescoring was needed to ensure research-level reliability. The researchers scored the same reports and discussed rubric scores. They then scored additional reports and repeated this process until an interrater reliability of at least 85% was obtained for all abilities.

Students were evaluated with rubrics on a total of 22 scientific subabilities during the mechanics portion of the school year. Table VIII contains the scientific subabilities that students were evaluated on at least three times. Each of these subabilities were displayed as a 100% stacked bar chart and we ran a Wilcoxon signed ranks test [87] on the scores to determine if whole class growth over the course of the semester was statistically significant. We then compared whole class growth for class sections that had the opportunity to revise and those who did not revise after receiving feedback.

TABLE VIII. Scientific subabilities frequently assessed in investigations. Note that lab 9 is omitted from analysis. Students designed their own experiments and chose the subabilities on which to evaluate their work. It is not possible to measure whole-class growth because so many different subabilities were selected.

| | Lab 1 | Lab 2 | Lab 3 | Lab 4 | Lab 5 | Lab 6 | Lab 7 | Lab 8 |
|---|---|---|---|---|---|---|---|---|
| Is able to identify sources of experimental uncertainty | X | X | | X | X | X | X | X |
| Is able to evaluate specifically how identified uncertainties may affect the data | X | X | | X | X | X | X | X |
| Is able to identify assumptions made in using the mathematical procedure | X | X | | | | X | X | X |
| Is able to determine specifically the way in which assumptions might affect the results | X | X | | | | X | X | X |
| Is able to record and represent data in a meaningful way | X | X | | X | | X | X | |
| Is able to make a reasonable prediction based on the hypothesis | | X | X | | | X | X | |
| Is able to make a reasonable judgment about the hypothesis | | X | X | | | X | X | |
| Is able to make a judgment about the results of the experiment | | | X | | | X | X | |
| Is able to evaluate the results by means of an independent method | | X | | | X | | | X |

*Individual student contributions.* We conducted a quantitative analysis of the revision history of each Google document to track individual student contributions on laboratory reports. The benefit of using teacher-provided Google documents is that we had access to the revision history. Each document was run through an analysis program that rebuilt the document and kept track of individual keystrokes (character count). A quantitative analysis of each lab report revealed the number of additions and deletions made by each student for each investigation; this includes revisions. This analysis did not account for whether the students worked collaboratively on a single device as the program was only able to determine which account typed each keystroke. Table IX contains an example of what this output looks like for a single laboratory investigation.

As a part of this analysis, it is possible to obtain information about which students contributed to which scientific subabilities. We performed a case study analysis to answer the research question of how individual students contribute when writing up and revising laboratory reports in teams. Although we have data from ten sections of honors physics, the findings shared in this paper are limited to a single class section from instructor 1. Selecting one class was purposeful so all students worked on their investigations under the same conditions (amount of class time, teacher directions, etc.). After conducting a whole class analysis, the work of two students was chosen to showcase the step-by-step process in the classroom.

The first student contributed to many scientific subabilities on the initial written submission and participated in the revision process. The second student contributed minimally to both the initial written submission and the revisions, and their participation was limited to the same scientific subabilities.

### 2. Surveys

Surveys were administered at the start and end of each school year.

*E-CLASS.* We administered the Colorado Learning Attitudes about Science Survey for Experimental Physics (E-CLASS) survey to all sampled students during the second and third years of data collection. This survey assesses student views and attitudes when doing experiments in the classroom and has a score range from $-30$ to $+30$ [84]. It was not possible for the high school students to complete the survey on the digital platform monitored by the University of Colorado, Boulder because they were minors at the time of instruction. As a result, we conducted our own statistical analysis of the data following the procedure used by Wilcox and Lewandowski [88]. The 5-point Likert item scale was collapsed to a 3-point Likert item scale. We combined the responses of “(dis)agree” and “strongly (dis)agree” into a single category as the difference between “(dis)agree and strongly (dis)agree” is not necessarily consistent across students [89]. This new

TABLE IX. Sample character count student contribution data for a single lab investigation.

| Name | Total added | Total deleted | % total contribution | Final version | % of final version | % of revisions |
|---|---|---|---|---|---|---|
| Jack | 2994 | 719 | 27.24 | 2754 | 35.96 | 33.68 |
| Scott | 1973 | 1483 | 25.35 | 1010 | 13.19 | 16.05 |
| Divya | 4142 | 672 | 35.31 | 2385 | 31.14 | 33.78 |
| Arpun | 1537 | 113 | 12.10 | 1510 | 19.72 | 16.50 |

3-point scale was coded as favorable (+1), neutral (0), or unfavorable (−1) with respect to expert response. The sum of each individual's score was obtained. We ran general descriptives for central tendency and variability on the dataset. Overall student scores were graphed and the averages and standard deviations found. Finally, we ran a Wilcoxon signed ranks test to determine if the difference in pre- and postscores was statistically significant.

*Scientific abilities confidence survey.* We ran descriptive statistics on the pre- and postresponses to the scientific ability survey. Once again, the Wilcoxon signed ranks test was run to test for significance in change in confidence response pre and post.

#### 3. Relationship between artifacts

We ran a bivariate correlation to determine if there were any significant relationships between different individual student artifacts including average contribution, the number of times they revised, FCI gain, E-CLASS difference, and scientific ability confidence gain. A Pearson correlation coefficient was used for continuous data and a Spearman correlation for all other cases.

## III. FINDINGS

In this section we present findings related to the effects of the opportunity to revise written laboratory reports on student scientific ability development. A table summarizing the laboratory reports by year and instructor is available in the Appendix (Table XIV). We also present pre- and postsurvey results for the E-CLASS and scientific abilities confidence surveys as well as relationships between the opportunity to revise and other individual student artifacts.

### A. Scientific abilities

We explored changes to whole class overall mastery of scientific abilities and how the opportunity to revise affected student growth by studying subabilities. In the findings we present an overview of this growth for three types of subabilities: uncertainty-associated abilities, assumption-associated abilities, and hypothetico-deductive reasoning. These subabilities were selected for this paper because students consistently had multiple exposures over all three years of data collection. Their work related to these particular subabilities clearly addresses how the opportunity to revise affects student growth. Findings are presented from year 3 for uncertainty-associated abilities because both Instructors provided multiple opportunities for engagement with these abilities. We present findings for all three years for assumption-associated abilities and hypothetico-deductive reasoning as there were only two formal exposures (student work evaluated with the rubrics) to each during year 3.

#### 1. Uncertainties

The first formal evaluation with rubrics of experimental uncertainties in year 3 was on an investigation two to three weeks into the school year. During the first few days of school, students were informally introduced to uncertainty associated abilities through smaller classroom activities. These activities required students to use uncertainty associated abilities but their work was not evaluated with the rubrics. In this investigation, students were asked to use their mathematical model for constant motion to predict where two cars would meet when released a certain distance apart. The students needed to express the predicted meeting point as a range and identify the sources of uncertainty. The students in instructor 1's class wrote lab report 1 collaboratively in teams of four while the students in instructor 2's class wrote individual reports. This is the only report that students in instructor 2's class wrote individually; the remaining reports were all written collaboratively (as discussed in Sec. II C 1). Students in both classes struggled with identifying sources of experimental uncertainty on lab 1 (Fig. 5)—their focus was on either instrumental or random uncertainty, not both. Over the course of the semester, students in both sections improved at identifying all sources of experimental uncertainty. This reached saturation for instructor 1's students after lab 2 and varied slightly as the cognitive demand of each task increased. While students in instructor 2's class also improved at identifying sources of experimental uncertainty, they did not reach saturation by the end of the semester.

Furthermore, students in both classes demonstrated significant proficiency (a score of 2 or 3) on their initial

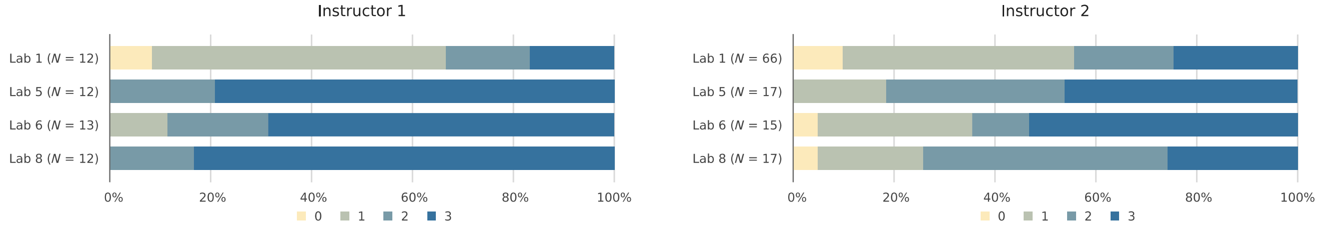


FIG. 5. "Is able to identify sources of experimental uncertainty" rubric scores for year 3. The bars represent the percentage of groups whose initial laboratory reports received the scores shown on the rubrics. 0, missing; 1, inadequate; 2, needs improvement; and 3, adequate.

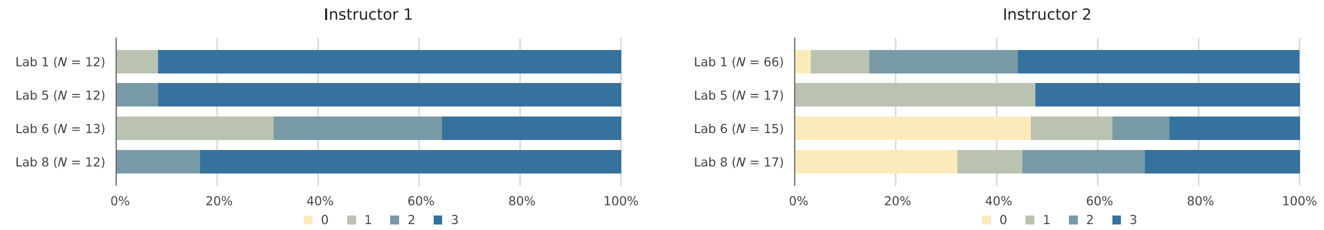


FIG. 6. “Is able to evaluate specifically how identified experimental uncertainties may affect the data” rubric scores for year 3. The bars represent the percentage of groups whose initial laboratory reports received the scores shown on the rubrics. 0, missing; 1, inadequate; 2, needs improvement; and 3, adequate.

submission with evaluating the identified experimental uncertainties (Fig. 6). Instructor 1’s students’ scores reached saturation during lab 5 and decreased significantly on lab 6 ($p < 0.001$) (where the task was more cognitively demanding) before returning to saturation on lab 8. However, after the first investigation, the majority of instructor 2’s students struggled with correctly evaluating the uncertainties and stating the final value as a range of possible outcomes. After lab 1, their demonstrated proficiency remained at or below 50% for the remainder of the semester.

Students in instructor 1’s class had the opportunity to revise and resubmit their work. Figure 7 shows how revisions affected the initial submission for uncertainty-associated abilities. Although there appears to be fluctuation in scores throughout the semester, both the content and cognitive demand of the investigations increased with time (Table V). Significant changes in scores occurred for both subabilities when new content was introduced (i.e., impulse and momentum in lab 6). Groups returned to saturated demonstrated proficiency on lab 8.

### 2. *Assumptions*

In year 1, students had five formal opportunities to work with assumption associated abilities. As with experimental uncertainties, students were introduced to assumptions during informal classroom activities. Figure 8 shows how revisions affected the initial submission for assumption associated abilities. Consistent with uncertainty associated abilities, student scores fluctuated throughout the semester with each lab when they had the opportunity to revise.

In year 2 we saw a similar pattern in student growth with these subabilities over the course of the semester (Fig. 9). Their ability to identify assumptions made in using the mathematical procedure was at saturation for lab 1 (a less cognitively demanding task). This decreased to 50% on lab 2 before returning to close to 100% for the remaining lab reports. The ability to determine specifically the way in which assumptions might affect the results was more challenging for students. On most investigations, the majority of student groups stated the effects of their assumptions but did not attempt to validate any of them (the difference between a score of 2 and a score of 3). There was also a drop from 80% to 30% between lab 1 and lab 2 that is consistent with other fluctuations in the course.

In year 3, students only had two formal opportunities to demonstrate proficiency with assumption associated abilities. As seen in Figs. 10 and 11, this was not enough for growth. Students in both classes started the year with more than 80% proficiency identifying assumptions on lab 1. Over the course of the semester, proficiency on both of these abilities remained constant. Instructor 2’s students hovered around 70% proficiency for the ability to determine specifically the way in which assumptions might

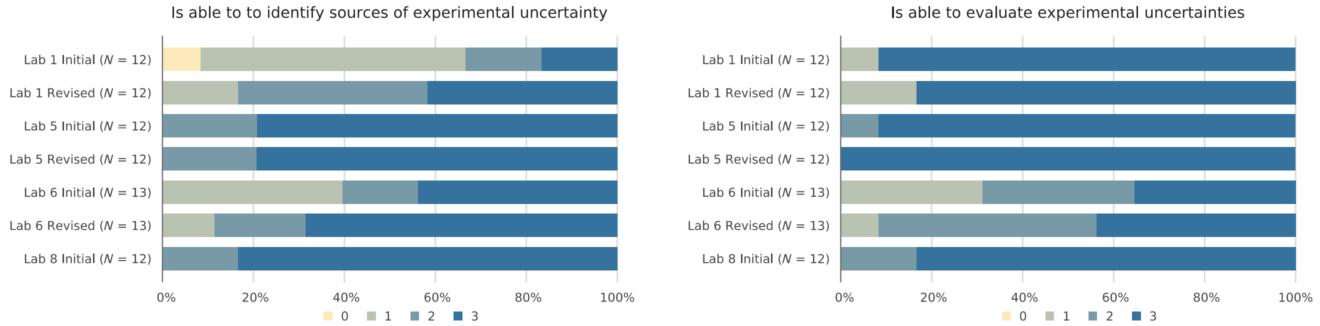


FIG. 7. Initial rubric scores for year 3 instructor 1 uncertainty associated abilities for all groups. Revised rubric scores for all groups who chose to revise. Uncertainty associated abilities for lab 8 were only scored by researchers so students did not have an opportunity to revise. The bars represent the percentage of groups whose laboratory reports received the scores shown on the rubrics. 0, missing; 1, inadequate; 2, needs improvement; and 3, adequate.

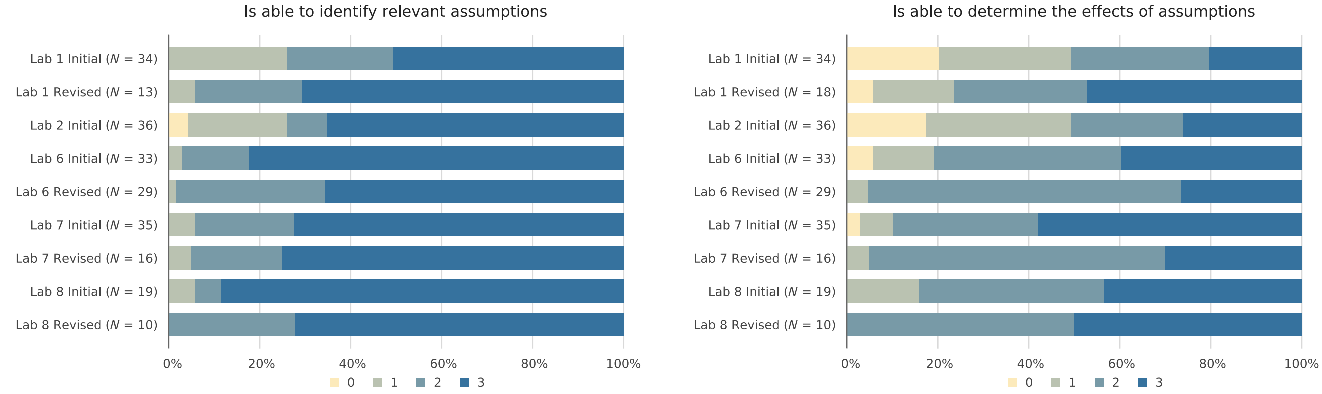


FIG. 8. Initial rubric scores for year 1 assumption associated abilities for all groups. Revised rubric scores for all groups who chose to revise. Assumption associated abilities for lab 2 were only scored by researchers so students did not have an opportunity to revise. The bars represent the percentage of groups whose submissions received each rubric score. 0, missing; 1, inadequate; 2, needs improvement; and 3, adequate.

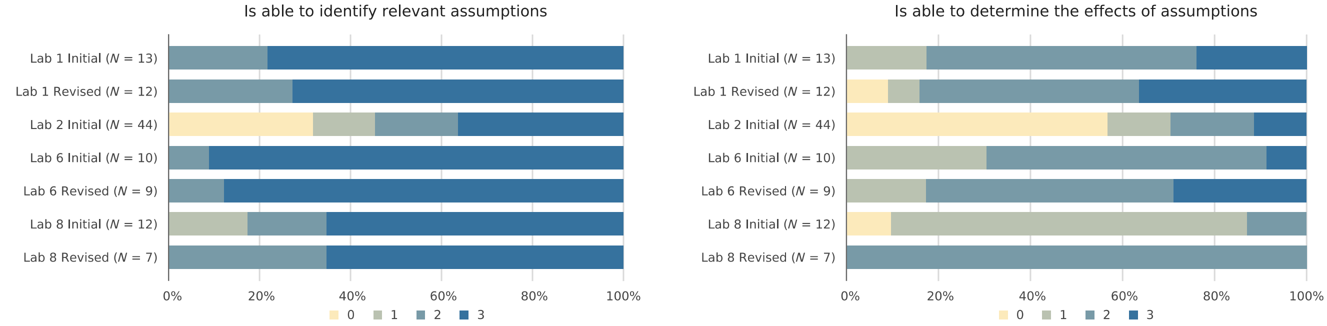


FIG. 9. Initial rubric scores for year 2 assumption associated abilities for all groups. Revised rubric scores for all groups who chose to revise. Assumption associated abilities for lab 2 were only scored by researchers so students did not have an opportunity to revise. The bars represent the percentage of groups whose submissions received each rubric score. 0, missing; 1, inadequate; 2, needs improvement; and 3, adequate.

affect the results while instructor 1's students (who had the opportunity to revise) remained at 60% proficiency.

### *3. Hypothetico-deductive reasoning*

Hypothetico-deductive reasoning abilities are known to be challenging for students [12]. In year 1, students had three formal opportunities where they were asked to make a reasonable prediction of the outcome of a testing experiment using the hypothesis under test. They also completed numerous informal activities in class that engaged them in this reasoning. As shown in Fig. 12, the opportunity to revise did not impact students' demonstrated proficiency on their initial submission (constant rubric scores). This differs from uncertainty and assumption associated abilities where the opportunity to revise contributed to increased proficiency.

During year 3, students in both sections completed multiple informal ISLE investigations and two formal lab investigations where they were asked to make a reasonable prediction of an outcome of a testing experiment using the hypothesis under test. As seen in Fig. 13, students in instructor 1's class were able to demonstrate over 60% proficiency on the first lab but the entire class proficiency

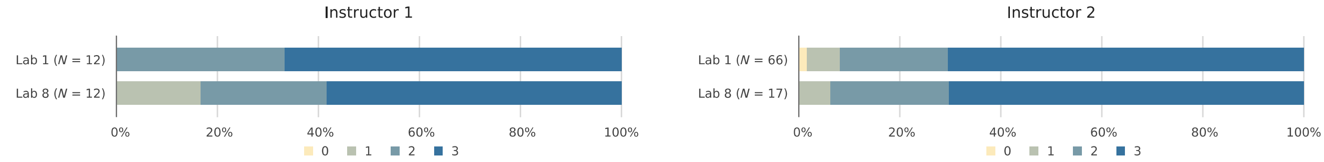


FIG. 10. "Is able to identify assumptions made in using the mathematical procedure" rubric scores for year 3. The bars represent the percentage of groups or individual students (lab 1, instructor 2) whose initial submission received each score on the rubric. 0, missing; 1, inadequate; 2, needs improvement; and 3, adequate.

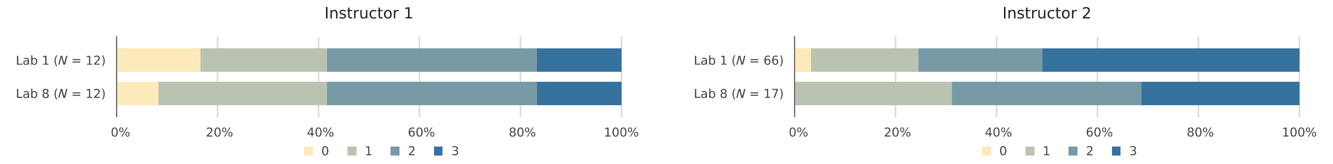


FIG. 11. “Is able to determine specifically the way in which assumptions might affect the results” rubric scores for year 3. The bars represent the percentage of groups or individual students (lab 1, instructor 2) whose initial submission received each score on the rubric. 0, missing; 1, inadequate; 2, needs improvement; and 3, adequate.

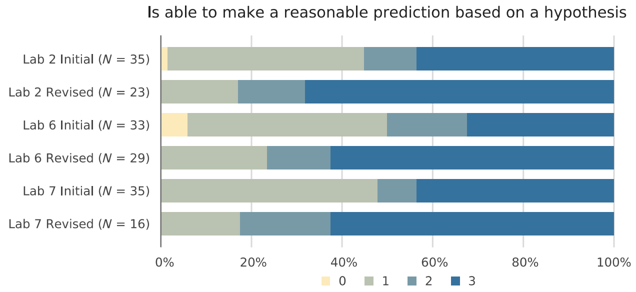


FIG. 12. Initial rubric scores for year 1 the ability to make a reasonable prediction based off the hypothesis under test for all groups. Revised rubric scores for all groups who chose to revise. The bars represent the percentage of groups whose submissions received each rubric score. 0, missing; 1, inadequate; 2, needs improvement; and 3, adequate.

dropped to 35% on lab 6. For instructor 2’s class the majority of students were unable to make a reasonable prediction based on the hypothesis and their demonstrated proficiency remained roughly constant (lab 3 average rubric score, $0.548 \pm 0.850$; lab 6 average rubric score, $0.629 \pm 0.794$) throughout the semester.

## B. Survey results

### *1. E-CLASS survey*

Students in year 2 and year 3 took the Colorado Learning Attitudes about Science Survey for Experimental Physics (E-CLASS) at the beginning and end of the school year. 137 out of the 161 enrolled students completed both the pre- and post-survey. Table X contains the average score per student.

The difference between the pre- and post-test means is statistically insignificant regardless of year or instructor (Table X). Many students entered the class with positive, expertlike beliefs in their attitudes toward experimental physics and this remains after a year of ISLE approach high school physics. In traditional lab courses at the college level, student attitudes decrease significantly over the course of the semester [18]. Question 17 (when I encounter difficulties in the lab, my first step is to ask an expert, like the instructor) is the only question where their responses did not correlate with those of experts.

Although the pre-post means for E-CLASS scores was statistically insignificant, students showed a statistically significant change in attitude on some individual questions. In year 2, students demonstrated a positive change ($p < 0.005$) in their understanding of the purpose of doing physics experiments (question 16). In year 3, students in both classes demonstrated a positive increase ($p < 0.001$ for instructor 1 and $p < 0.05$ for instructor 2) in their attitude toward assumptions (question 3).

### *2. Scientific abilities confidence survey*

81% of the 231 students enrolled in instructor 1 and instructor 2’s sections of honors physics completed the pre- and postscientific ability confidence survey. Consistent with the E-CLASS scores, student mean responses to each question started high. Instructor 1’s students demonstrated statistically significant increase in self-reported confidence on five of the ten questions. Instructor 2’s students demonstrated statistically significant increase in self-reported confidence on nine of the ten questions. This is inconsistent with their development of scientific abilities. The only survey item that did not show statistically significant change for any grouping of participants was “applying constructed knowledge to the real world.” Table XI contains significance for year 3 student responses

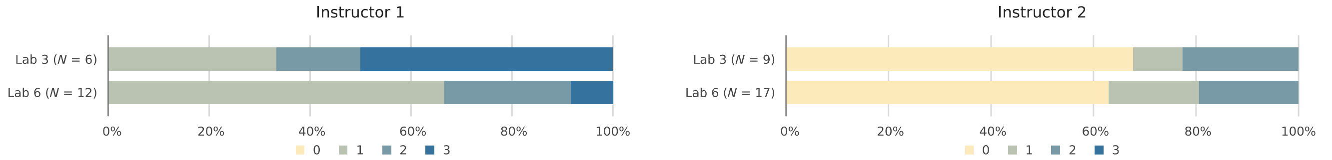


FIG. 13. Initial group scores for year 3 instructor 1 and instructor 2 “is able to make a reasonable prediction based on a hypothesis.” Students in instructor 1’s sections had the opportunity to revise. The bars represent the percentage of groups whose laboratory reports received the scores shown on the rubrics. 0, missing; 1, inadequate; 2, needs improvement; and 3, adequate.

TABLE X. E-CLASS pretest and post-test means and signed ranks tests for significance for different subgroups of the sample population.

| Student group | Pretest mean | Post-test mean | Significance |
|---|---|---|---|
| All students ($N = 137$) | $15.63 \pm 5.20$ | $14.94 \pm 6.76$ | $p = 0.247$ |
| Year 2 $N = 40$) | $14.85 \pm 5.84$ | $13.58 \pm 6.76$ | $p = 0.135$ |
| Year 3 Inst. 1 ($N = 41$) | $16.54 \pm 4.83$ | $17.10 \pm 6.47$ | $p = 0.397$ |
| Year 3 Inst. 2 ($N = 56$) | $15.52 \pm 4.97$ | $14.34 \pm 6.76$ | $p = 0.262$ |

TABLE XI. Scientific ability survey responses per year. Asterisks show confidence gains at significance level $^{*}p < 0.05$ and $^{**}p < 0.001$.

| Ability | Year 3 instructor 1 ($N = 42$) | Year 3 instructor 2 ($N = 57$) |
|---|---|---|
| Using multiple representations to articulate your reasoning | 0.004* | 0.003* |
| Applying constructed knowledge to the real world | 0.147 | 0.140 |
| Differentiating between a hypothesis and a prediction | 0.143 | 0.004* |
| Writing a prediction based on a hypothesis | 0.254 | 0.002* |
| Analyzing the outcome of an experiment and taking into account assumptions | 0.010* | <0.001** |
| Analyzing the outcome of an experiment and taking into account experimental uncertainties | 0.145 | 0.001** |
| Conducting an observational experiment | 0.016* | 0.002* |
| Conducting a testing experiment | 0.060 | 0.001** |
| Conducting an application experiment | 0.001** | 0.000** |
| Posing your own research question | 0.002* | 0.039* |

when only the students of instructor 1 had the opportunity to revise.

### C. Relationship between artifacts

We ran a bivariate correlation to explore relationships between individual student contributions and responses on different artifacts. Table XII contains the outputs for these relationships for all students. There are statistically significant relationships between average contribution and number of times students revise, their FCI gain, E-CLASS score, and self-reported confidence with the scientific abilities.

### D. Individual student contributions

This paper presents preliminary findings from a larger case study analysis that investigates how individual students contribute when writing up lab reports collaboratively. For consistency in instruction and delivery, we choose to focus these preliminary findings on one class section (22 total students) from year 2 of data collection for students who had the opportunity to revise. Year 2 was selected because the students wrote the majority of reports collaboratively and they were evaluated multiple times on many scientific abilities (leading to repeat exposure). Students wrote five group reports and two individual reports during the mechanics portion of the semester.

TABLE XII. Correlation outputs across all artifacts with independent contributions for all students in the study. Asterisks show correlations at significance level $^{*}p < 0.05$ and $^{**}p < 0.001$.

| | Average contribution | Number of revisions | FCI gain | E-CLASS difference | Scientific ability confidence gain |
|---|---|---|---|---|---|
| Average contribution | 1.000 | 0.606** | 0.145* | 0.016 | 0.037 |
| Number of revisions | 0.606** | 1.000 | 0.163* | 0.246* | 0.187 |
| FCI gain | 0.145* | 0.163* | 1.000 | 0.021 | 0.120 |
| E-CLASS difference | 0.016 | 0.246* | 0.021 | 1.000 | 0.273* |
| Scientific ability confidence gain | 0.037 | 0.187 | 0.120 | 0.273* | 1.000 |

We tallied the number of times students made significant contributions to the laboratory reports and revisions. For the purpose of this study, significant contributions are defined as >20% as students worked in teams of four. Ninety five percent of students contributed to the majority (3 out of 5) of group reports with 17 (77%) writing more than 20% of at least one report. Five students (23%) contributed more than 20% of the written work on all five reports. Seventeen students (77%) revised at least one investigation. Seven (32%) students participated in revising at least three reports, and two students revised all five of their lab reports. The students who contributed most frequently engaged with the largest variety of scientific abilities, regardless of grouping. Those with minimal contributions typically wrote the same sections of each report.

Some of these students were like Riley. Riley contributed more than 20% of all written work on three of the five reports and always revised based on teacher feedback. They exhibited a high FCI gain (0.65) and a positive change in expertlike attitude toward experimental physics on the E-CLASS survey (11 → 17). On written reports they frequently

**Author(s):** Casey (32.638 %), Riley (42.246 %), Rowan (9.309 %), Harper (15.807 %), Instructor (0.000 %)*, **UNKNOWN (0.000 %)**, * - does not contribute to relative percentages

Hypothesis: The relationship, F avg*±t=m±v , is true.

Prediction: If the ±v is 0.45m/s, the mass of the system is 0.675 kg, ±t is 0.97s, and the track is frictionless, then the average force exerted on the dynamics cart by the pulley's hanging weight should equal 0.4019N ±0.062N

Calculations:
50g weight:
F avg ±t=m±v
F avg (0.97s±0.019s)=(0.675kg±0.135kg)(0.45m/s±0.009m/s)
F avg =(0.30375kgm/s±0.00122kgm/s)±(0.97s±0.019s)
F avg =0.313N±0.0626N

100g weight:
F avg ±t=m±v
F avg (0.61s±0.0122s)=(0.712kg±0.145kg)(0.69m/s±0.0137m/s)
F avg =(0.49128kgm/s±0.00199kgm/s)±(0.61s±0.0122s)
F avg =0.8177N±0.163N

Analysis:
In our experiment we acknowledged the presence of multiple experimental uncertainties, mostly instrumental uncertainties because all of the data was collected using instruments. We did not have any random uncertainties because we only did one trial of each situation to collect our data.

Judgement:
Looking the data collected and taking into account the assumptions and uncertainties, we judge our hypothesis to hold true. The range of average force calculated with the data we collected differed from the average force given by the motion detector through LoggerPro by 0.0263. Though the value from LoggerPro did not match ours, it is close enough to show that our hypothesis is consistent, since we made assumptions that we were unable to take into account. These assumptions included that the track was frictionless, and that the track was on a flat surface. We also assumed the cart accelerated constantly on the track. These factors all affect the outcome of our calculations, especially if the track was not completely flat, which also depends on surface of the table. A unsmooth surface can cause our data to be slightly off, for the car will not accelerate at a constant rate. Looking at the difference in values between our calculations and LoggerPro, we can account for them in our assumptions; hence, we judge our hypothesis to be consistent.

FIG. 14. Excerpts for student contributions to lab 6. Colors are attributed to student keystrokes.

completed different tasks including the procedure, data tables, experimental uncertainty analysis, discussion of assumptions, and the discussion. Riley also inserted images of diagrams, drawings, and mathematical models on several investigations. Riley's written contributions to lab 8 were minimal as another classmate took control of writing up the final report, despite the strengths of their peers.

Alternatively, several students in the class contributed more like Rowan. Rowan's work always accounted for less than 10% of the final product. They only contributed to revisions on two reports. Their FCI gain was also high (0.52) and they exhibited a positive change in expertlike attitude toward experimental physics on the E-CLASS survey (13 → 15). On written reports, they frequently completed the same tasks including the data table and experimental uncertainty identification and calculations.

Figure 14 contains an excerpt from a written report for lab 6 when Riley and Rowan were in the same group. In this investigation, students were tasked with designing an experiment to test the impulse-momentum relationship. Riley (dark green) contributed to the prediction and much of the discussion. Rowan (pink) contributed primarily to calculations and uncertainty analysis. Blue text belongs to other students in the group.

## IV. DISCUSSION

### A. Research question No. 1

*How does the opportunity to revise written laboratory reports affect whole class development of scientific abilities?* We used a microgenetic approach to examine student written laboratory reports at multiple points throughout the mechanics portion of the high school curriculum. In the classroom of instructor 1, students always had the opportunity to revise and resubmit their work after receiving feedback. In the classroom of instructor 2, students only received feedback on their work. Consistent with other ISLE approach scientific ability development studies [12,71,90], the development of scientific abilities was not linear and students' demonstrated proficiency varied greatly with the cognitive difficulty of the task, number of repeat exposures to each ability, and the opportunity to revise. In addition to the opportunity to revise, our findings show how the role of the instructor and individual student contributions play a role in student performance in addition to this formative assessment strategy.

#### *1. The opportunity to revise*

The opportunity to revise work after receiving teacher feedback is an integral part of the ISLE approach [31,33] and aided students in instructor 1's class in the development of uncertainty and assumption-associated abilities. It also provides opportunities for students to develop a growth mindset [36,37] while learning how to fail productively [91–94]. As evidenced in the findings, student teams in instructor 1's class during year 3 showed greater gains than those in instructor 2's classes. Over 68% of students and student groups elected to revise their reports in year 1, 50% in year 2, and 75% in year 3. This means on average, 75% of teams in year 3 received twice as many exposures to the scientific abilities than their peers in instructor 2's classroom.

Research has shown that our brains are capable of reshaping themselves as learning and emotion produce physical changes in the brain [95,96]. In ISLE approach classrooms, providing multiple opportunities for engagement with the scientific abilities gives students an opportunity to strengthen connections between neurons. The frequent firing of neurons in the cortex creates and reinforces synapses in the network [95]. Studies have also shown that practice can increase brain density associated with a particular task [97].

The students in instructor 1's classroom had an opportunity to receive teacher feedback and then work together to improve upon their current level of understanding. Peer instruction benefits growth and understanding [98,99]. In this low-stakes environment, students are able to fill in the gaps in their understanding and strengthen neural network connections associated with thinking like a physicist.

Figures 5–11 illustrate that students have greater gains on those subabilities with repeat exposures compared to those with only two or three. Students were able to demonstrate proficiency with uncertainty and assumption-associated abilities, but struggled with making a reasonable prediction based on a hypothesis. This variation in proficiency for different subabilities is consistent with findings at the college level [12,100]. For example, when asked to make a reasonable prediction based on a hypothesis, instructor 1's high school students remained at approximately 50% proficiency throughout the year, despite increasing their demonstrated proficiency on individual investigations to approximately 80% after teacher feedback and revisions.

College students in prior studies had four [12] and nine [100] formal opportunities to practice making predictions. While not the same level of proficiency as uncertainty and assumption-associated abilities, they were able to demonstrate growth over the course of the semester. This means that more formal exposures and opportunities to learn from their mistakes are beneficial for challenging scientific abilities. Revisions are one type of repeat exposure that may help high school students develop this reasoning skill as they provide additional opportunities to strengthen synapses in the brain.

As shown in Table VII, not all students elected to revise their work. When asked about the revision process at the end of the school year, some students commented that they did not need to complete revisions to improve their grade or understanding. These students felt that the number of classroom exposures was sufficient for them to maintain strong neural connections related to the scientific abilities. These students consistently earned 2's or 3's on the rubrics. Others acknowledged that they were too busy with other commitments and did not have the time to revisit their work. One student remarked how they elected not to revise at first, but as the year progressed, recognized the benefits.

> At first, I didn't take advantage of the opportunity to revise. I didn't have time to do more work, and was happy with my grade. However, as the content became more complicated, I realized that the skills I struggled with in earlier labs were actually important for my learning and revisions were helpful in better understanding the classroom activities.

As the learning is integrated into labs, it became important for this student to learn from their mistakes in order to better understand the topics being studied. Treating the investigations the same as other classroom activities provided students with a consistent learning routine; they knew there would be an opportunity to try again if they did not understand the topic during their first exposure. Several other students commented that they appreciated this formative assessment opportunity because it removed the pressure that comes with grades and allowed them to demonstrate understanding at their own pace. These students found themselves engaging with the productive failure routines encouraged by the ISLE-approach classroom practices implemented by instructor 1. These feelings are consistent with the goals of formative assessment opportunities in the classroom [41–44], support for productive failure and mistake approaches to instruction [91,101,102], and the second intentionality of ISLE (enhancing student well being) [33]. Furthermore, resilience with challenging tasks will be beneficial to these learners in their future studies and careers [103].

### *2. The role of the instructor*

One difference between the high school and college level courses is that the laboratory component is fully integrated into the instruction at the high school level and there is one instructor who leads all types of the activities. As a result, the students in this study were consistently exposed to the scientific abilities through other classroom activities and interacted with their instructor on a more consistent basis. They received more feedback [39] regarding their performance than those at the college level.

As can be seen in the findings, in some instances the opportunity to revise did not play a role in student demonstrated proficiency. Instructor 2's students in year 3 for the ability to "determine specifically the way in which assumptions might affect the results" performed higher than instructor 1's students, who had the opportunity to revise. After conversations with instructor 2, there are several possible explanations for this difference. First, instructor 2 provided more guidance and/or scaffolding when circulating the room during the investigations. While instructor 1 also provided support during the investigations, increased temporary support structures can help students resolve cognitive dissonance when developing laboratory skills and practices with new concepts [104]. Second, instructor 2 summarized the experimental goals and key attributes of the investigation with the students before the final report was due. Known as "time for telling," [105] this sharing of "normative" information provides students an opportunity to reflect on their own ideas and compare them to those of practicing physicists. Students may have revised their reports prior to submission. Third, while both instructors routinely incorporated these subabilities into instruction, it is possible that instructor 2 placed greater emphasis on discussing the effects of assumptions in the classroom in their whole-class discussions. Further investigation is needed into informal exposures and the influence of the high school instructor on student scientific ability development.

### *3. Individual student growth*

Although our research question explores the effect of revisions on whole class development of scientific abilities, it is important to understand how individual students contribute to collaboratively written lab reports. Although ISLE classrooms strive for equitable learning practice [33] it was relatively unknown how students engaged in collaborative writing of laboratory reports. Based on the qualitative analysis of student collaborative reports written on school provided Google accounts, we found that individual student contributions varied based on the cognitive demand of the task and grouping of students. We learned that the written work is not always distributed equally among students.

This work is part of a broader research study to learn how individual students contribute to group lab reports. From the preliminary case study findings, out of a class of 22 students, some (like Riley) frequently completed the majority of the writing and those students had an opportunity to practice many different scientific abilities both on the initial and revised reports. Other students, such as Rowan, consistently contributed to the same scientific abilities and did not participate in the revision process. Students who took ownership of their learning took advantage of formative assessment opportunities [106] and due to their resubmission had additional exposure to the scientific abilities. Awareness of which abilities students contribute to on a consistent basis may help teachers develop norms and expectations to assist students in more well-rounded development of these practices.

Based on these findings, it is imperative that instructors are mindful of student personalities so as not to marginalize learners. To assist with dominating personalities, instructors can assign rotating roles or require students to focus on different subabilities for each investigation. Classroom teachers need to be aware of these trends in order to better structure groupwork for more balanced student participation and student identity development [107–111].

### B. Research question No. 2

*What is the effect of scientific ability development in an ISLE-approach classroom on student attitudes toward and confidence with experimental physics?*

#### *1. Attitudes*

All sections sampled for this study started the school year with high expertlike attitudes toward experimental physics. Many of the students enrolled in this course have an interest in science and plan to pursue careers in STEM fields, which may explain their initial E-CLASS results. Their attitudes over the course of the year remained positive. These findings are consistent with studies done at the college level for students who entered the classroom already embodying expertlike beliefs [112]. Students such as Riley who contributed to a variety of scientific subabilities on the written reports and revised demonstrated more expertlike attitudes toward experimental physics on the pre- and postsurvey. Other students, such as Rowan, who contributed to the same scientific subabilities had more varied scores on the pre- and postsurvey. Some reported positive growth while others reported a decrease in expertlike attitude. This variation in attitude toward experimental physics correlates directly to student engagement in the revision of written lab reports (Table XII). We can conclude that these repeat exposures have a direct influence on student attitudes because these learners are physically reshaping their brains when they reflect on their understanding and work to improve. When students feel successful upon persevering through a challenging task, these positive emotions [95] might reinforce their neural connections associated with scientific ability development.

Although student attitudes did not show statistically significant change over the course of the school year, there is a positive relationship between some attitudes and some scientific abilities. The only E-CLASS survey question that directly correlated to one of the scientific abilities was question 3 (when doing a physics experiment I don't think much about assumptions). Students showed significant growth with assumption-associated abilities and their increase in expert-like attitude toward assumptions was statistically significant. This means that not only did students think more about assumptions in the laboratory, but they were also able to productively apply this reasoning to their investigations. Perhaps this positive correlation exists because students were frequently exposed to assumption-associated abilities in the classroom, allowing for deep learning in the cerebral cortex [113]. Alternatively, as evidenced in the findings, not all students were proficient at identifying and explaining the effects of assumptions on their initial lab report submission. Reflecting upon their assumptions and making changes to demonstrate improved understanding may have led to an increased awareness of the importance of considering assumptions when designing and performing an experiment.

#### *2. Confidence*

Furthermore, there is a direct correlation between student attitudes and student self-reported confidence with the scientific abilities (Table XII). Helping students develop self-regulatory capabilities is an important element of teaching [38]. Students who felt more confident about engaging with the scientific abilities also held positive, expertlike attitudes toward experimental physics. Unlike student attitudes, student self-reported confidence increased significantly over the course of the school year for all groupings of students. While students in instructor 1's class showed gains in both their self-reported confidence and the scientific abilities over the three-year period, students in instructor 2's class during year 3 showed the greatest confidence gains on the survey. This was surprising because these students did not demonstrate as much growth with the scientific abilities as their peers in instructor 1's classes. This discrepancy is consistent with the findings of Kruger and Dunning who showed that students have difficulty assessing their own learning [114].

In this study, this discrepancy may be due to several factors. First, instructor 1 identified as female and instructor 2 identified as male. Student perceptions of the two instructors as teachers based on their genders may have influenced their confidence [115]. Second, the two instructors may have provided different levels of emphasis in the classroom with respect to the varying scientific abilities. Perhaps instructor 2 spent more time explicitly defining the abilities and their importance when conducting an investigation. Students think they learn better from teachers who provide direction instruction [116,117]. The students in instructor 2's class may have interpreted listening to prelab directions and postlab summaries as learning how to apply them to new situations. Therefore, these students may have perceived the difficulty of the investigations as simpler than they were [118]. Finally, all lab reports were rescored by researchers. Instructor 2's original scoring was higher than that of instructor 1's. Therefore, students in instructor 2's class consistently received feedback that the work they produced was in line with the expectations of the scientific abilities. This positive feedback may have had a direct influence on their confidence in the lab [39]. This provides an opportunity to further explore the effects of teacher feedback in ISLE approach classrooms.

### C. Implications for instruction

Although this study was conducted with a specific population of students, there are several implications for instruction for any student population that can be drawn from the findings. National organizations promote learning science through experimentation and reasoning [5–9]. The results of this study are grounded in multiple theoretical frameworks and decades of research on how people learn. The ISLE approach focuses on student development of habits of mind that will benefit them in the future as well as feeling good about themselves as physicists [31,33].

The ISLE approach successfully engages students in the practices of scientists and does not cause a drastic shift in student attitude toward experimental physics. The opportunity to revise work, without penalty, provides students with additional exposures to the scientific abilities and helps them develop proficiency in many of these abilities. To help students develop proficiency, educators need to make the epistemological structure of the development of ideas transparent for students. As seen in the findings, this process is not linear and takes time; one or two exposures to each scientific ability is not enough. Classroom instructors need to provide support for students to develop the resources necessary to productively engage with the scientific abilities. Educators need to appropriately scaffold the learning so challenging moments help students grow and find success.

### D. Limitations

There are several limitations to this study that need to be considered when drawing conclusions about the findings. First, this study was conducted in a single school with two teachers implementing the ISLE approach. We did not examine student development of scientific abilities in a non-ISLE setting. Furthermore, this study was completed based on artifacts from the classroom. All of the surveys were administered at the beginning and end of the school year so there could have been other factors, not considered here, that may have affected gains. Because of the fact that the participants were minors, we could not video record their interactions in the laboratory so can only base our analysis and findings on written artifacts.

Second, without being able to video record these interactions, having students submit their work digitally limited our ability to determine who contributed to each part of the lab report. Because the computer program attributed contributions to the Google account under which the student was logged in, there was no way to know if students worked collaboratively on the same machine(s) in the classroom or one student completed the majority of the writing. As a result, we are only able to speak to whole-class growth with the scientific abilities and whole-class correlations between the opportunity to revise, scientific ability development and attitudes regarding experimental physics.

The case studies provide us with some insight into individual student contributions, but their scientific subability scores are still tied to a variety of peers. Future investigations into the artifacts are necessary to understand individual student growth in a collaborative environment. Furthermore, the classroom instructor also made modifications to the labs themselves over the three-year period; this is most significant in year 3 when a second instructor was added.

With respect to student groups, changing them frequently could have helped limit the number of times one student completed the majority of the work. Additionally, learning to work with a variety of students with different strengths and personalities is a valuable skill all learners need to be successful in their future courses and careers [1,4].

Third, the students sampled for this study come from a high performing district and the written samples were collected from the honors section of the physics course. As a result, the learning progression for these students may not be generalizable to other populations. Although we do not aim to generalize the findings from this study, we believe there are aspects of the instruction, such as revisions, that may be applicable to other classrooms. This cannot be a conclusive statement until we examine how students in other classrooms at other high schools develop scientific abilities under the ISLE approach.

## ACKNOWLEDGMENTS

I am grateful to Eugenia Etkina for her advice and support during data collection, analysis, and the preparation of this paper. I thank the teachers for their support during data collection and Diane Jammula and Elana Resnick for reading this manuscript and providing feedback. I also thank the referees for reading this paper and providing feedback that strengthened the publication.

# APPENDIX

TABLE XIII. Scientific ability confidence survey administered by classroom teacher. Students were asked to rank their comfort and confidence with the following scientific abilities. 0 = I do not know what this is; 5 = I could teach someone how to do it.

| Ability | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| Using multiple representations to articulate your reasoning. | | | | | | |
| Applying constructed knowledge to real world situations. | | | | | | |
| Differentiating between a hypothesis and a prediction. | | | | | | |
| Writing a prediction based on a hypothesis. | | | | | | |
| Analyzing the outcome of an experiment and taking into account assumptions. | | | | | | |
| Analyzing the outcome of an experiment and taking into account experimental uncertainty. | | | | | | |
| Conducting an observational experiment. | | | | | | |
| Conducting a testing experiment. | | | | | | |
| Conducting an application experiment. | | | | | | |
| Posing your own research question. | | | | | | |

TABLE XIV. Summary of data collected and number of exposures to each ability category discussed in the findings; data exclude lab 9 from years 2 and 3 as students were able to select from all 22 subabilities.

| Year | Instructor | Revisions | Collaborative or individual | Uncertainty-associated abilities | Assumption-associated abilities | Hypothetico-deductive reasoning abilities |
|---|---|---|---|---|---|---|
| 1 | 1 | Yes | Each section wrote 2 individual lab reports and 5 collaborative lab reports (rotating) | 5 exposures | 5 exposures | 3 exposures |
| 2 | 1 | Yes | Each section wrote 1 individual lab report and 5 collaborative lab reports | 6 exposures | 4 exposures | 2 exposures |
| 3 | 1 | Yes | Each section wrote 6 collaborative lab reports | 4 exposures | 2 exposures | 2 exposures |
| 3 | 2 | No | Each section wrote 1 individual and 5 collaborative lab reports | 4 exposures | 2 exposures | 2 exposures |

[1] OECD, *Education at a Glance 2018: OECD Indicators* (OECD Publishing, Paris, 2018), 10.1787/eag-2018-en.

[2] L. McNeil and P. Heron, Preparing physics students for 21st-century careers, Phys. Today **70**, No. 11, 38 (2017).

[3] W. M. Roth and J. Désautels, Educating for citizenship: Reappraising the role of science education, Can. J. Sci. Math. Tech. Educ. **4**, 149 (2004).

[4] World Economic Forum, report on the future of jobs (2020). Retrieved from https://www3.weforum.org/docs/WEF_Future_of_Jobs_2020.pdf.

[5] European Commission, Directorate-General for Research and Innovation, *Science education for responsible citizenship—Report to the European Commission of the expert group on science education* (Publications Office, Luxembourg, 2015), https://data.europa.eu/doi/10.2777/12626.

[6] *America's Lab Report: Investigations in High School Science*, Tech. Rep., edited by S. R. Singer, M. L. Hilton, and H. A. Schweingruber (Committee on High School Science Laboratories: Role and Vision; National Research Council, Washington, DC, 2005).

[7] American Association of Physics Teachers, AAPT Recommendations for the Undergraduate Physics Laboratory Curriculum, Tech. Rep. (American Association of Physics Teachers, College Park, MD, 2014), https://www.aapt.org/Resources/.

[8] Next Generation Science Standards (2013), http://www.nextgenscience.org/.

[9] College AP Courses (2017). Retrieved from https://www.collegeboard.org.

[10] N. G. Holmes and H. J. Lewandowski, Investigating the landscape of physics laboratory instruction across North America, Phys. Rev. Phys. Educ. Res. **16**, 020162 (2020).

[11] E. Etkina, S. Murthy, and X. Zou, Using introductory labs to engage students in experimental design, Am. J. Phys. **74**, 979 (2006).

[12] E. Etkina, A. Karelina, and M. Ruibal-Villasenor, How long does it take? A study of student acquisition of scientific abilities, Phys. Rev. ST Phys. Educ. Res. **4**, 020108 (2008).

[13] N. G. Holmes and D. A. Bonn, Doing Science or Doing a Lab? Engaging Students with Scientific Reasoning during Physics Lab Experiments, presented at PER Conf. 2013, Portland, OR, 10.1119/perc.2013.pr.034.

[14] E. M. Smith and N. G. Holmes, Evaluating instructional labs' use of deliberate practice to teach critical thinking skills, Phys. Rev. Phys. Educ. Res. **16**, 020150 (2020).

[15] R. F. Lippmann, Students' understanding of measurement and uncertainty in the physics laboratory: Social construction, underlying concepts, and quantitative analysis. Ph.D. thesis, University of Maryland, 2003.

[16] P. G. Ellingsen and J. A. Stovneg, Student laboratory reports: An approach to improving feedback and quality, Eur. J. Phys. **39**, 035701 (2018).

[17] D. Hu and B. M. Zwickl, Examining students' views about validity of experiments: From introductory to Ph.D. students, Phys. Rev. Phys. Educ. Res. **14**, 010121 (2018).

[18] B. R. Wilcox and H. J. Lewandowski, Developing skills versus reinforcing concepts in physics labs: Insight from a survey of students' beliefs about experimental physics, Phys. Rev. Phys. Educ. Res. **13**, 010108 (2017).

[19] D. Buggé and E. Etkina, Reading between the lines: Lab reports help high school students develop abilities to identify and evaluate assumptions, presented at PER Conf. 2016, Sacramento, CA, 10.1119/perc.2016.pr.008.

[20] D. Buggé, The short and long-term effects of the ISLE approach on high school physics students' attitudes and development of science-process abilities. Ph.D. thesis, Rutgers University, 2020.

[21] M. Marušić and J. Slisko, Influence of three different methods of teaching physics on the gain in students' development of reasoning, Int. J. Sci. Educ. **34**, 301 (2012).

[22] A. L. Stephens and J. J. Clement, Documenting the use of expert scientific reasoning processes by high school physics students, Phys. Rev. ST Phys. Educ. Res. **6**, 020122 (2010).

[23] W. M. Roth, Experimenting in a constructivist high school physics laboratory, J. Res. Sci. Teach. **31**, 197 (1994).

[24] A. R. Carter, One hundred years later, introductory labs are poised for change, Phys. Teach. **59**, 97 (2021).

[25] C. Gosling, Authentic independent investigations in high school physics laboratories, Phys. Teach. **59**, 48 (2021).

[26] S. Laursen, Levers for change: An assessment of progress on changing stem instruction, Tech. Rep. (American Association for the Advancement of Science, Washington, DC, 2019).

[27] A. Szott, Open-ended laboratory investigations in a high school physics course: The difficulties and rewards of implementing inquiry-based learning in a physics lab, Phys. Teach. **52**, 17 (2014).

[28] E. Etkina, A. Karelina, M. Ruibal-Villasenor, D. Rosengrant, R. Jordan, and C. E. Hmelo-Silver, Design and reflection help students develop scientific abilities: Learning in introductory physics laboratories, J. Learn. Sci. **19**, 54 (2010).

[29] National Research Council, *A Framework for K–12 Science Education: Practices, Crosscutting Concepts, and Core Ideas* (National Academies Press, Washington, DC, 2012).

[30] A. Madsen, S. B. McKagan, and E. C. Sayre, How physics instruction impacts students' beliefs about learning physics: A meta-analysis of 24 studies, Phys. Rev. ST Phys. Educ. Res. **11**, 010115 (2015).

[31] E. Etkina, D. T. Brookes, and G. Planinsic, *Investigative Science Learning Environment: When Learning Physics Mirrors Doing Physics* (Morgan & Claypool Publishers, San Rafael, CA, 2019).

[32] E. Etkina, Millikan award lecture: Students of physics—Listeners, observers, or collaborative participants in physics scientific practices?, Am. J. Phys. **83**, 669 (2015).

[33] D. T. Brookes, E. Etkina, and G. Planinsic, ISLE: An epistemologically authentic approach to student-centered inquiry learning, Phys. Rev. Phys. Educ. Res. **16**, 020148 (2020).

[34] C. S. Dweck and E. L. Leggett, A social-cognitive approach to motivation and personality, Psychol. Rev. **95**, 256 (1988).

[35] H. Grant and C. S. Dweck, Clarifying achievement goals and their impact, J. Pers. Soc. Psychol. **85**, 541 (2003).

[36] D. C. McClelland, *Human Motivation* (Scott Foresman, Oakland, NJ, 1984).

[37] D. S. Yeager and C. D. Dweck, Mindsets that promote resilience: When students believe that personal characteristics can be developed, Educ. Psychol. **47**, 302 (2012).

[38] D. Carless, D. Salter, M. Yang, and J. Lam, Developing sustainable feedback practices, Stud. Higher Educ. **36**, 395 (2011).

[39] J. Hattie and H. Timperley, The power of feedback, Rev. Educ. Res. **77**, 81 (2007).

[40] D. R. Sadler, Formative assessment and the design of instructional systems, Instr. Sci. **18**, 119 (1989).

[41] P. Black and D. Wiliam, Assessment and classroom learning, Assess. Educ. Princ. Pol. Pract. **5**, 7 (1998).

[42] P. Black and D. Wiliam, Developing the theory of formative assessment, Educ. Assess. Eval. Acc. **21**, 5 (2009).

[43] I. Clark, Formative assessment: Assessment is for self-regulated learning, Educ. Psychol. Rev. **24**, 205 (2012).

[44] P. R. Pintrich, *The role of goal orientation in self-regulated learning*, in *Handbook of Self-Regulation* (Academic Press, New York, 2000), pp. 451–502.

[45] P. Cobb and J. Bowes, Cognitive and situated learning perspectives in theory and practice, Educ. Res. **28**, 4 (1999).

[46] E. Etkina, A. Van Heuvelen, S. White-Brahmia, D. T. Brookes, M. Gentile, S. Murthy, D. Rosengrant, and

A. Warren, Scientific abilities and their assessment, Phys. Rev. ST Phys. Educ. Res. **2**, 020103 (2006).
[47] J. G. Greeno, A. M. Collins, and L. B. Resnick, Cognition and learning, Handbook Educ. Psych. **77**, 15 (1996).
[48] D. Hestenes, M. Wells, and G. Swackhamer, Force Concept Inventory, Phys. Teach. **30**, 141 (1992).
[49] P. Black and D. Wiliam, The reliability of assessments, Assess. Learn. 119 (2006).
[50] D. L. Butler and P. H. Winnie, Feedback and self-regulated learning: A theoretical synthesis, Rev. Educ. Res. **65**, 245 (1995).
[51] National Research Council, *How People Learn: Bridging Research, and Practice* (National Academies Press, Washington, DC, 1999).
[52] D. Boud and E. Molloy, Rethinking models of feedback for learning: The challenge of design, Assess. Eval. Higher Educ. **38**, 698 (2013).
[53] C. Moskovitz and D. Kellogg, Inquiry-based writing in the laboratory course, Educ. Sci. **332**, 919 (2011).
[54] N. Lerner, Laboratory lessons for writing and science, Written Commun. **24**, 191 (2007).
[55] C. Haagen-Schuetzenhoefer, Improving the quality of lab reports by using them as lab instructions, Phys. Teach. **50**, 430 (2012).
[56] L. O. P. Rivard, A review of writing to learn in science: Implications for practice and research, J. Res. Sci. Teach. **31**, 969 (1994).
[57] V. Howard, Thinking on paper: A philosopher's look at writing, in *Varieties of Thinking: Essays from Harvard's Philosophy of Education Research Center*, edited by V. A. Howard (Routledge, New York, 1988), Vol. 84.
[58] J. A. Reynolds, C. Thaiss, W. Katkin, and R. J. Thompson, Writing-to-learn in undergraduate science education: A community-based, conceptually driven approach, CBE Life Sci. Educ. **11**, 17 (2012).
[59] B. Hand, C. W. Wallace, and E. M. Yang, Using a science writing heuristic to enhance learning outcomes from laboratory activities in seventh-grade science: Quantitative and qualitative aspects, Int. J. Sci. Educ. **26**, 131 (2004).
[60] C. W. Keys, B. Hand, V. Prain, and S. Collins, Using the science writing heuristic as a tool for learning from laboratory investigations in secondary science, J. Res. Sci. Teach. **36**, 1065 (1999).
[61] J. T. Stanley and H. J. Lewandowski, Lab notebooks as scientific communication: Investigating development from undergraduate courses to graduate research, Phys. Rev. Phys. Educ. Res. **12**, 020129 (2016).
[62] M. Eblen-Zayas, Comparing electronic and traditional Lab Notebooks in the advanced lab, in *Laboratory Instruction: Beyond the First Year* (American Association of Physics Teachers, College Park, MD, 2015), pp. 28–31.
[63] W. B. Lane, Letters home as an alternative to lab reports, Phys. Teach. **52**, 397 (2014).
[64] J. R. Hoehn and H. J. Lewandowski, Framework of goals for writing in physics lab classes, Phys. Rev. Phys. Educ. Res. **16**, 010125 (2020).
[65] J. Fitzgerald, Research on revision in writing, Rev. Educ. Res. **57**, 481 (1987).
[66] C. Wagner and A. Vaterlaus, Promoting formative assessment in high school teaching of physics, Latin Am. J. Phys. Educ. **6**, 410 (2012).
[67] E. Cruz, H. Dias, and G. Kortemeyerlll, The effect of formative assessment in Brazilian university physics courses, Rev. Bras. Ensino Fis. **33**, 4315 (2011).
[68] R. J. Dufresne and W. J. Gerace, Assessing-to-learn: Formative assessment in physics instruction, Phys. Teach. **42**, 428 (2004).
[69] C. Plybour, Integrating formative assessment into physics instruction: The effect of formative vs summative assessment on student physics learning and attitudes, Ph.D. thesis, Western Michigan University, 2015.
[70] M. Zhu, O. L. Liu, and H. Lee, The effect of automated feedback on revision behavior and learning gains in formative assessment of scientific argument writing, Comp. Educ. **143**, 103668 (2020).
[71] E. Etkina, A. Karelina, S. Murthy, and M. Ruibal-Villasenor, Using action research to improve learning and formative assessment to conduct research, Phys. Rev. ST Phys. Educ. Res. **5**, 010109 (2009).
[72] https://sites.google.com/site/scientificabilities/.
[73] H. Andrade and A. Valtcheva, Promoting learning and achievement through self-assessment, Theory Into Practice **48**, 12 (2009).
[74] E. Panadero, J. Alonso-Tapia, and E. Reche, Rubrics vs self-assessment scripts effect on self-regulation, performance and self-efficacy in pre-service teachers, Stud. Educ. Eval. **39**, 125 (2013).
[75] W. J. Popham, What's wrong-and what's right-with rubrics, Educ. Leader. **55**, 72 (1997).
[76] E. Dawson, Social justice and out-of-school science learning: Exploring equity in science television, science clubs and maker spaces, Sci. Educ. **101**, 539 (2017).
[77] Y. Malini Reddy and H. Andrade, A review of rubric use in higher education, Assess. Eval. Higher Educ. **35**, 435 (2010).
[78] J. W. Creswell, *Research Design—Qualitative, Quantitative, and Mixed Methods Approaches* (Sage Publications, Inc., Los Angeles, CA, 2013).
[79] https://pum.islephysics.net.
[80] E. Etkina, D. Brookes, G. Planinsic, and A. Van Heuvelen, *Active Learning Guide for College Physics: Explore and Apply* (Pearson, San Francisco, CA, 2019), 2nd ed.
[81] M. B. Miles and A. M. Huberman, *Qualitative Data Analysis a Methods* (Sage Publications, Inc., Thousand Oaks, CA, 1994).
[82] R. R. Hake, Interactive-engagement versus traditional methods: A six-thousand student survey of mechanics test data for introductory physics courses, Am. J. Phys. **66**, 64 (1998).
[83] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevPhysEducRes.19.020166 for sample laboratory investigation.
[84] 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).

[85] R. S. Siegler and K. Crowley, The microgenetic method: A direct means for studying cognitive development., Am. Psych. **46**, 606 (1991).

[86] G. Calais, Microgenetic analysis of learning: Measuring change as it occurs, Natl. Forum Appl. Educ. Res. J. **21**, 1 (2008).

[87] F. Wilcoxon, Individual comparisons by ranking methods, Biom. Bull. **1**, 80 (1945).

[88] 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).

[89] M. Lovelace and P. Brickman, Best practices for measuring students' attitudes toward learning science, CBE Life Sci. Educ. **12**, 606 (2013).

[90] S. Faletic and G. Planinsic, How the introduction of self-assessment rubrics helped students and teachers in a project laboratory course, Phys. Rev. Phys. Educ. Res. **16**, 020136 (2020).

[91] M. Kapur, Productive failure, Cognit. Instr. **26**, 379 (2008).

[92] M. M. Clifford, Thoughts on a theory of constructive failure, Edu. Psychol. **19**, 108 (1984).

[93] R. A. Schmidt and R. A. Bjork, New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training, Psychol. Sci. **3**, 207 (1992).

[94] M. Kapur and K. Bielaczyc, Designing for productive failure, J. Learn. Sci. **21**, 45 (2012).

[95] J. E. Zull, *The Art of Changing the Brain: Enriching the Practice of Teaching by Exploring the Biology of Learning* (Routledge, Philadelphia, PA, 2002).

[96] S. Dehaene, *How We Learn: Why Brains Learn Better Than Any Machine…For Now* (Penguin, Westminster, London, 2021).

[97] B. Draganski, C. Gaser, V. Busch, G. Schuierer, U. Bogdhan, and A. May, Changes in grey matter induced by training, Nature (London) **427**, 311 (2004).

[98] C. Turpen and N. D. Finkelstein, The construction of different classroom norms during Peer Instruction: Students perceive differences, Phys. Rev. ST Phys. Educ. Res. **6**, 020123 (2010).

[99] J. Watkins and E. Mazur, Retaining students in science, technology, engineering, and mathematics (STEM) majors, J. Coll. Sci. Teach. **42**, 36 (2013).

[100] D. Buggé, J. Rutberg, S. Ahmed, R. Zisk, and D. Jammula, Development of hypothetico-deductive skills in an ISLE-based lab taught by novice instructors, Phys. Educ. **58**, 035013 (2023).

[101] J. Kerrigan, K. Weber, and C. Chinn, Effective collaboration in the productive failure process, J. Math. Behav. **63**, 100892 (2021).

[102] N. Steenhof, N. N. Woods, and M. Mylopoulos, Exploring why we learn from productive failure: Insights from the cognitive and learning sciences, Adv. Health Sci. Educ. **25**, 1099 (2020).

[103] J. D. Bransford and D. L. Schwartz, Rethinking transfer: A simple proposal with multiple implications, Rev. Res. Educ. **24**, 61 (1999).

[104] F. L. Braca and C. S. Kalman, Comparison of laboratorials and traditional labs: The impacts of instructional scaffolding on the student experience and conceptual understanding, Phys. Rev. Phys. Educ. Res. **17**, 010131 (2021).

[105] D. L. Schwartz and J. D. Bransford, A time for telling, Cognit. Instr. **16**, 475 (1998).

[106] S. Brookhart, C. Moss, and B. Long, Promoting student ownership of learning through high-impact formative assessment practices, J. Multi. Disc. Eval. **6**, 52 (2009).

[107] N. Shah, J. A. Christensen, N. A. Ortiz, A. K. Nguyen, S. Byun, D. Stroupe, and D. L. Reinholz, Racial hierarchy and masculine space: Participatory in/equity in computational physics classrooms, Comp. Sci. Educ. **30**, 254 (2020).

[108] Z. Chen, R. B. Lawson, L. R. Gordon *et al.*, Groupthink: Deciding with the leader and the devil, Psychol. Rec. **46**, 581 (1996).

[109] P. Aggarwal and C. L. O'Brien, Social loafing on group projects: Structural antecedents and effect on student satisfaction, J. Market. Educ. **30**, 255 (2008).

[110] K. N. Quinn, M. M. Kelley, K. L. McGill, E. M. Smith, Z. Whipps, and N. G. Holmes, Group roles in unstructured labs show inequitable gender divide, Phys. Rev. Phys. Educ. Res. **16**, 010129 (2020).

[111] E. M. Stump, M. Dew, S. Jeon, and N. G. Holmes, Taking on a manager role can support women's physics lab identity development, Phys. Rev. Phys. Educ. Res. **19**, 010107 (2023).

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

[113] D. A. Kolb, *Experiential Learning* (Prentice Hall, Englewood Cliffs, NJ, 1984).

[114] J. Kruger and D. Dunning, Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments, J. Pers. Soc. Psychol. **77**, 1121 (1999).

[115] G. Potvin, Z. Hazari, R. H. Tai, and P. M. Sadler, Unraveling bias from student evaluations of their high school science teachers, Sci. Educ. **93**, 827 (2009).

[116] D. Watkins, Student perceptions of factors influencing tertiary learning, Higher Educ. Res. Dev. **3**, 33 (1984).

[117] L. Deslauriers, L. S. McCarty, K. Miller, K. Callaghan, and G. Kestin, Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom, Proc. Natl. Acad. Sci. U.S.A. **116**, 19251 (2019).

[118] K. A. Burson, R. P. Larrick, and J. Klayman, Skilled or unskilled, but still unaware of it: How perceptions of difficulty drive miscalibration in relative comparisons, J. Pers. Soc. Psychol. **90**, 60 (2006).