

Assessing student engagement with teamwork in an online, large-enrollment course-based undergraduate research experience in physics

Alexandra Werth^{1,2,*} Kristin Oliver^{1,2} Colin G. West,¹ and H. J. Lewandowski^{1,2}

¹*Department of Physics, University of Colorado, Boulder, Colorado 80309, USA*

²*JILA, National Institute of Standards and Technology and University of Colorado, Boulder, Colorado 80309, USA*



(Received 14 April 2022; accepted 27 September 2022; published 25 October 2022)

[This paper is part of the Focused Collection on Instructional labs: Improving traditions and new directions.] Over the last decade, course-based undergraduate research experiences (CUREs) have been recognized as a way to improve undergraduate science, technology, engineering, and mathematics education by engaging students in authentic research practices. One of these authentic practices is participating in teamwork and collaboration, which is increasingly considered to be an important component of undergraduate research experiences and laboratory classes. For example, the American Association of Physics Teachers Recommendations for the Undergraduate Physics Laboratory Curriculum suggest that one of the goals for students in physics labs should be to develop “interpersonal communication skills” through “teamwork and collaboration.” Teamwork can have tremendous benefits for students, including increased motivation, creativity, and reflection; however, it can also pose an array of new social and environmental challenges, such as differing styles of communication, levels of commitment, and understanding of concepts. It can also be difficult for lab course instructors to evaluate and assess. In this work, we study student teamwork in a large-enrollment physics CURE. The CURE was specifically designed to emphasize teamwork as a scientific practice. We use the two sources of data, the adaptive instrument for regulation of emotions questionnaire and students’ written memos to future researchers, to measure the students’ teamwork goals, challenges, self, co-, and socially shared regulations, and perceived goal attainment. We find that students overwhelmingly achieved their teamwork goals by overcoming obstacles using primarily socially shared regulatory strategies, and that the vast majority of students felt teamwork was an essential part of their research experience. We discuss implications for the design of future CUREs and lab courses and for lab instructors desiring to assess teamwork in their own courses.

DOI: [10.1103/PhysRevPhysEducRes.18.020128](https://doi.org/10.1103/PhysRevPhysEducRes.18.020128)

I. INTRODUCTION

During the Fall 2020 semester, the introductory physics lab for engineers and physical science majors at University of Colorado (CU) Boulder was redesigned due to the COVID-19 pandemic. To continue to serve the students’ needs in the changing educational context, we developed and deployed a course-based undergraduate research experience (CURE) [1]. A CURE is a formal course in which all students can enroll and engages the entire class of students in an authentic scientific research project, seeking to answer a question that is of genuine interest to the scientific community [2]. In this manner, a CURE seeks to replicate

much of what is present in “traditional” undergraduate research experiences (UREs), which have been shown to have numerous positive benefits [2–9], including development of more expertlike epistemology [9,10] and increased persistence in STEM [11]. However, traditional research experiences can come with institutional and systemic barriers to entry [2,12–15], many of which are generally absent in a CURE, since by definition it is an authentic research experience open to all who enroll. A CURE can also reach large numbers of students at once; while several very large CUREs exist in fields like biology [16,17], to our knowledge our course represented the first instance of a large-enrollment (> 400 students per semester) CURE in the field of physics. Students in our course analyzed the energy distribution of hundreds of individual solar flares to collectively address the infamous “coronal heating problem” [18] (for more details on the course and CURE research question, see Sec. II C).

By their nature, CUREs also typically feature authentic scientific research practices, like collaboration and peer review [3,9]. To reinforce on this, our course was designed

*alexandra.werth@colorado.edu

Published by the American Physical Society under the terms of the *Creative Commons Attribution 4.0 International license*. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

around three explicit goals: teaching essential experimental research skills, providing a positive and motivating encounter with experimental physics, and fostering productive teamwork [1]. Of the three, the degrees of success for the first and second can be readily measured: the first through traditional classroom assessments, and the second through affective or attitudinal research instruments [19,20]. But while direct studies of teamwork in physics labs do exist [21,22], attempts to measure teamwork outcomes appear not to be as common as measurements of student gains with respect to content, skills, or affect. Even when teamwork is an explicit goal of the course, a typical measurement strategy may simply be an informal survey of the students' self-reported teamwork experiences [23].

Teamwork, however, is increasingly viewed as an essential feature of a lab course. For example, The American Association of Physics Teachers (AAPT) recommendations for the undergraduate physics laboratory curriculum [24] suggests that students in physics labs should develop “interpersonal communication skills” through “teamwork and collaboration.” After all, teamwork in a classroom environment can have tremendous benefits for students, such as increasing their motivation, creativity, and self-reflection [25]. Moreover, since learning “skills” (as contrasted with physics content) is often a major feature of lab courses [26], and since collaboration and teamwork is an integral part of modern scientific practice, it is logical for a lab course to try specifically to foster positive and productive teamwork experiences for students. Indeed, with research and business becoming increasingly global endeavors, many have argued that it is now more important than ever for students to be prepared to collaborate with diverse groups of people, including online if needed [27,28].

Hence, it seems valuable to include learning goals related to teamwork when designing or redesigning a lab course, and attainment of those learning goals should be measurable. At the same time, teamwork and collaborative learning pose significant challenges for students even in traditional classroom conditions [29,30], not to mention the new challenges posed if the course takes place online [31,32]. If teamwork is a goal for a course, then ideally we should attempt to measure the extent to which it is achieved, as well as the challenges that impede that goal and the strategies students use to overcome them.

Broadly, this paper provides an in-depth example of an attempt to qualitatively and quantitatively assess the success of a teamwork goal, and the associated challenges faced by the students, in a large-enrollment physics CURE. In doing so, we highlight some of the relevant forms of data that can be collected to assess teamwork methods, not merely in a CURE, but in any lab course, or perhaps even in a traditional undergraduate research experience.

More specifically, our results address the following questions in the context of our CURE course:

RQ1. What were some of the goals that students had for teamwork?

RQ2. What were some of the challenges that students faced during teamwork?

RQ3. What regulations did students use to overcome teamwork challenges? In particular, how much did students rely on self-regulated learning (SRL), co-regulated learning (CoRL), and socially shared regulated learning (SSRL) strategies?

RQ4. To what extent did students achieve their goals through teamwork?

We answer these questions by examining two primary data sources: (i) a final reflection written by students from the Fall 2020 semester in the form of a memo to future researchers and (ii) closed responses to the adaptive instrument for regulation of emotions (AIRE) questionnaire [33] from the Spring 2021 semester, a research-based tool which measures students' teamwork goals, challenges, their use of self, co-, and shared regulation to address challenges, and their perceived goal attainment.

Given the inherent challenges of group work and the additional barriers to effective collaboration that have been documented in remote courses during the pandemic [34], we expected that, despite our clear objective to the contrary, teamwork might nevertheless prove to be a primary point of frustration for the students. However, we found quite the opposite. In an initial analysis of student survey data from the Fall 2020 semester of the CURE course, we found that students had overwhelmingly positive teamwork experiences, with over 80% of the class reporting that teamwork helped them stay motivated in the course, learn, and more successfully conduct their research [1]. In this work, we argue that our more detailed analysis of the four questions described above also gives substantial support to the conclusions that students had positive teamwork experiences in our course, and that they made appropriate use of various regulation strategies when challenges arose.

Finally, we provide a discussion theorizing about which elements of the course were responsible for the successful teamwork experiences, and how successful teamwork experience may influence the students in the future. For example, did a specific course element help foster more socially shared regulations as a response to the teamwork challenges students faced or was it the very nature of participating in a CURE? Was teamwork viewed positively by students because of the COVID-19 pandemic, or despite it? And together, did this teamwork experience help students build identity or a sense of belonging to the scientific community? As a result, we argue there is a need for even more quantitative methods in this area, for example, to evaluate student learning of specific teamwork skills, or the impact of teamwork experiences on a student's understanding of scientific identity and the nature of science itself.

This paper begins by providing the relevant background on CUREs (generally, as well as our specific course), student teamwork, and collaborative learning (Sec. II). It is

then organized as follows: In Sec. III, we discuss the two data sources (the memos to future researchers and the AIRE survey), as well as our analysis methods and the differences between the Fall 2020 and Spring 2021 semesters of the course. Section IV contains the results and discussions answering each of our four research questions and speculating about the remaining questions regarding teamwork in this course. Lastly, we conclude with implications for future instruction and research in Sec. V.

II. BACKGROUND

A. Course-based undergraduate research experiences

The past decade has seen significant work to both study the impact of CUREs and to codify their key features [3]. For example, a CURE should feature the implementation of real scientific practices by building on, and contributing to, current scientific knowledge, including traits like iteration within the scientific method, as well as collaboration and teamwork [3,9]. Perhaps the most important feature of a CURE is that it engages students in authentic discovery, in which the outcome of an investigation is initially unknown to all.

Currently, the majority of CUREs described in the literature are centered in chemistry or biology [35,36]. These CUREs often take place in upper-division courses [35], and are typically fairly small (< 100 students) [37], though there has also been work [16] to transform some large, introductory labs into CUREs. At present, there are few reported instances of CUREs in physics, and those that exist [38,39] are also typically small [40] or have not necessarily featured fully authentic discovery [41]. By contrast, the introductory physics course described in this work involved more than 400 students per semester, performing work intended to be genuinely publishable (the resulting work is currently in preparation [42]). To our knowledge, this represents the first instance of a large-scale CURE in an introductory physics lab, and provides an example of fostering genuine collaboration and teamwork despite significant challenges of scale.

B. Teamwork in labs

Small-group work has been shown to promote greater academic achievement, increase favorable attitudes toward learning, and improve persistence in STEM courses [43]. However, social learning situations are typically more challenging to navigate than independent learning situations because students need to overcome challenges that emerge due to the social nature of the tasks in addition to the regular challenges faced in any learning setting [29,30,33]. For example, students participating in group work face challenges such as irreconcilable personal goals [44,45], differing styles of work and communication [46], differing levels of commitment, concentration, or

standards [47], and differing levels of prior knowledge, understanding of concepts, or power [48].

While teamwork is emphasized as an essential skill in the sciences and engineering, evaluation of teamwork using quantitative assessment tools has been comparatively rare, particularly in physics education research. However, there is a relevant research-based survey, AIRE questionnaire, which can be used to help assess teamwork in a quantitative manner and is built upon the concept of regulated learning theory [33].

The idea of regulated learning theory is that students constantly engage in cyclical processes of setting goals and strategies while learning, implementing strategies, monitoring their performance, and then reflecting and adapting [49]. This single cycle is often referred to as self-regulated learning theory as it involves only the individual learner. Subsequent theories (co-regulated learning theory and socially shared regulated learning theory [50]) have added complexity and nuance to this cycle that better reflect collaborative learning environments.

Co-regulation of learning can be thought of as “scaffolded” guidance, given from a more-able to a less-able individual [33]. Järvenoja, Volet, and Järvelä define co-regulation of learning to be “individuals’ various attempts to affect each other’s motivation, emotional state, cognitive actions, etc., for their own purpose or others’ benefit [33].”

In addition, entire groups of learners may mutually engage in socially shared regulation of learning, “where several individuals regulate their collective activity in a genuinely shared way [33].” For example, in a physics lab, students must take into consideration each other’s goals and define group goals (i.e., working toward shared goals that the group decides to pursue) [50].

In the same manner, the challenges students face during the “implementing strategies” stage of the learning cycle may be individual in nature (e.g., “I am having trouble understanding what my group is doing”) or collective (e.g., “We have trouble using this technology as a team”). And the regulations of these challenges may be individual or socially shared as well (e.g., “I need to change my goals to better match the goals of the rest of the group” versus “we need to work together to determine a more effective strategy to communicate”).

One example of using this theoretical model to evaluate team dynamics in an laboratory setting is recent work by Goñi *et al.* in the DILAB School of Engineering at Pontificia Universidad Católica de Chile [22]. This study uses the socially regulated learning theory (via the AIRE questionnaire [33]) to understand if face-to-face and online team dynamics differ concerning the prevalence of personal goals, team challenges, and individual or social strategies [22]. They found that both modalities report mostly the same prevalence of goals, challenges, and strategies [22]. These results contradict previous research that reported less student satisfaction [51] and more communication

challenges [52] in online teamwork. However, although they saw equivalency of teamwork in the two modalities (online and face to face), they noticed a trend in which students reported fewer task conflicts and team distraction challenges in the online setting, as well as more personal flexibility [22]. Although this may seem positive, they could be indicative of potential conflict avoidance that could be detrimental to the learning process, as task conflict is a known predictor of creativity and performance in team projects [22].

C. Course context and development

The CU CURE is a 15-week course in which students engage in a solar-physics research project called The Colorado PHysics Laboratory Academic Research Effort (C-PhLaRE). The project seeks to answer a longstanding question in solar physics by exploring a proposed mechanism responsible for heating the Sun's corona, which is millions of kelvin hotter than the photosphere, despite being much further out in space [18]. Some solar scientists speculate that the coronal heating is caused by many small “nanoflares” occurring constantly on the Sun, while others predict that the heating is dominated by a magnetohydrodynamic effects like Alfvén waves. The C-PhLaRE is working to determine if nanoflares are the dominate heating mechanism by calculating the solar flare frequency distribution (flare frequency rate versus energy in the long x-ray region) [53–55] on log-log scale from over one-thousand flare x-ray light curves and looking at the slope. A slope below -2 would imply that small flares are likely to dominate the heating while a slope greater than -2 would suggest magnetohydrodynamic waves are the primary cause [53].

In determining the energy of many different solar flares, each flare typically requires individualized attention: beginning and end points need to be chosen and appropriate baseline correction needs to be applied before the total energy can be integrated. Because of these individualize decisions, the problem is resistant to automated analysis by computer algorithms, which would reject many candidate events as unsuitable for analysis [56]. By contrast, these decisions are relatively straightforward for humans when they can be done on a case-by-case basis.

Using introductory-level physics and calculus, and with basic python data analysis tools that students learn during the course, small teams of students can calculate the total energy of an individual flare. Over three semesters, we had well over a thousand students working on the project, and collectively they analyzed hundreds of distinct flares. This large dataset gives a clear picture of the flare frequency distribution and directly addresses the question of the solar heating mechanism (results from the work are currently in preparation).

The teams performing this analysis typically consisted of three or four students. At the beginning of the course, students were surveyed about factors like their comfort

using video on Zoom, prior coding experience, coding confidence, time as a student at CU, declared major, gender identity, and time zone, which were collectively used to construct the teams. Students were assigned to their teams beginning in the second or third full week of the class (this differed slightly from semester to semester), and continued to work together consistently throughout the semester (except in cases of students withdrawing from the course, schedule changes, etc).

Since it was an explicit objective of the course that students should have productive and enjoyable teamwork experiences, we were intentional about the team assignments. In particular, we carefully avoided forming teams in which only one student did not identify as a man. This choice was motivated by literature on group dynamics showing that women sometimes have lower performance and poorer social cohesion in male-dominated group settings [57] and that a gendered division of roles can arise in “unstructured” lab environments such as ours [58]. Beyond this restriction, we sought to minimize the range of self-reported coding confidence within a team, to avoid situations where more experienced students would dominate the teamwork experience. These factors alone forced most of the team assignments, but other characteristics were used as needed; for example, if several students in a section were on the east coast.

Students worked together in these teams throughout the semester, which was divided into six phases:

1. Project on-boarding
2. Research plan development
3. Data analysis
4. Peer review
5. Calculating the slope
6. Documentation and reflection

In this paper, we highlight the first week of the project on-boarding phase, which consisted of explicit teamwork training. However, it is important to note that teamwork was required almost every week and students engaged in regular meta-cognitive reflections on their teamwork experiences. In addition, there were numerous instructional strategies that were used to emphasize teamwork throughout the course including, but not limited to the following:

1. Consistent and purposeful messaging about the importance of teamwork in the syllabus, lecture videos, and class assignments.
2. Pre-lab lectures (15–20 min online, asynchronous videos that students were required to watch prior to lab) presented examples of teamwork as a scientific practice giving examples of a large collaboration like CERN to smaller ones such as the lab the instructor runs at CU Boulder.
3. Explicit advice on “roles” students may take on in order to effectively code in Google Colab as a team.
4. Training for TAs on how to intervene if they see common teamwork challenges.

5. Incorporating authentic collaboratory research practices, such as peer review and a whole “group” meeting with the principal investigator.

Therefore, it is clear that no single element of the class can be said to be the exclusive cause of the teamwork gains seen in this study. More information detailing the entire course transformation can be found in an earlier publication [1].

In the teamwork training [59] during the first week of class, students were split into breakout rooms on Zoom to discuss a particular teamwork “scenario” based on the kinds of challenges that often arise when engaging in group work (note that, some challenges focused specifically on the context of a remote collaboration and others were more general). For example, one scenario was written from the perspective of a student whose more experienced teammate gradually took a larger and larger share of the work. Student groups then presented their scenario to the full lab section or another team and their “solutions,” such that students were exposed to multiple scenarios and teamwork strategies were written up and turned in as an assignment (more details on the training scenarios are available in Ref. [1]).

We note that there were some minor differences in the structure of the teamwork training between the Fall 2020 and Spring 2021 semester. In the Fall 2020 semester, students conducted the training in randomly assigned Zoom break-out rooms and received their team assignments the following week. However, in Spring 2021, students were assigned to their teams prior to the start of lab and conducted the teamwork training with their semester-long teams. In addition, the teamwork training discussion was shorted by 30 min during the Spring 2021 semester, so that students could spend that time practicing sharing their screens and working on the Colab notebooks collaboratively, which we had observed to be a common obstacle in the prior semester.

III. METHODOLOGY

Our data analysis began by analyzing student survey data from the Fall 2020 term of the CURE course, where we found over 80% of the class reporting that teamwork helped them stay motivated in the course, learn, and more successfully conduct their research [1]. In order to better understand these results and describe what happened during students’ teamwork experiences to lead to such positive outcomes, we turned to the “memo to future researchers,” written by students at the end of the Fall 2020 term (Sec. III A 1). The vast majority of students discussed teamwork in their memos and often described challenges they faced, regulations to overcome those challenges, and goals that could be attained through teamwork. To analyze these memos, we first investigated literature in this area and found the adaptive instrument for regulation of emotions questionnaire. From this prior work, we developed an *a priori* code book based on the

theoretical framework of socially shared regulation of learning and the items of the (AIRE) questionnaire [33]. This code book was used to analyze the memo data (Sec. III C 1). Given the strength of the code book, derived from the AIRE questionnaire, to represent the students’ written responses in the Fall 2020 term, we decided to administer the questionnaire itself at the end of the Spring 2021 semester (Secs. III A 2 and III C 2). In this work, we draw our conclusions from both the Fall 2020 student memos and the Spring 2021 AIRE questionnaire responses. Although the two data sources are from consecutive, different semesters, the results will not be presented in chronological order—instead, we will present the quantitative findings from the AIRE survey and then use the memos to demonstrate students teamwork experiences in the course, as well as highlight aspects of online teamwork that were not captured by the close-response survey. However, given that the data were collected from two different semesters (Sec. III B) we point out that the qualitative data analysis does not represent the voice of the same students who supplied AIRE responses in Spring 2021. A further discussion of this limitation as well as others can be found in Sec. III D.

A. Data collection

1. Memo to future researchers

In the Fall 2020 semester, since we were planning to continue the CURE research into the following semester, we asked the current students to write a one-page memo introducing the new undergraduate researchers to the project. This memo served both as an opportunity for metacognitive reflection and as a chance for students to engage in the authentic research practice of “handing off their experiment to the next researcher,” since scientific research projects are frequently not completed with the work of a single student.

In the memo assignment, which took place in the final week of the course, the students were asked to

1. Describe the project in terms such that someone new can understand what we are doing. What are the research goals? How did we achieve them?
2. Summarize what conclusions we can draw from our research results.
3. Discuss your personal experience (e.g., What did you learn this semester? Do you have any advice you would give them about the analysis, working with Colab, or teamwork?)
4. Suggest ideas for how this project could continue in the future. What should a new student be thinking about? What should their goals for the research be?

As shown in the prompt, teamwork was specifically mentioned as an example of what to describe about their personal experience, but it was not something that students were required to discuss. A large majority of the students (405/440) in the Fall 2020 semester completed the memo.

2. The AIRE survey

The AIRE survey, developed by Järvenoja, Volet, and Järvelä, is an instrument that adaptively measures SRL, CoRL, and SSRL within the context of student-led collaborative learning situations [33]. The theoretical foundation of the survey lies in contemporary theories of socially regulated learning (see Sec. II B), particularly focusing on regulation of motivations and emotions. The questions themselves were formed based on previous empirical work by the survey creators on motivation in real-life learning contexts [33,60], emotional experience in collaborative learning [33,61], and strategies for handling socially challenging learning activities [33,62]. The survey is broken into four sections: (i) Personal goals, (ii) socio-emotional challenges, (iii) an “adaptive” section on regulation of emotions, and (iv) perceived goal attainment.

The first section elicits students’ specific, personal goals for their current group learning activity (in our case, we modified the survey to ask about the entire teamwork experience in the PHYS-1140 CURE), and last section prompts students to reflect on the achievement of these personal goals indicated in part one. The idea is to capture some personal preferences that students bring to the activity, since these are assumed to have an impact on how students regulate their collaborative experience [33], this is consistent with SRL theories that believe the regulation processes are needed to constantly reflect on the achievement of personal goals or possibly change those goals in order to achieve success [33].

In the first section, students are given a list of goals [33] and asked to rank their importance on a 5-point Likert scale from “not at all important” to “extremely important” [63]. Next, students were asked to “Please select which of the goals above was most important to you?” and indicate “Which of the above was least important to you?” from a drop-down list [64].

The second section measures teamwork challenges faced by the students. This is based in the theory that in a social learning situation there are many distractions and challenges that can interfere with students achieving their personal goals [33]. In the survey, students are presented list of situations (i.e., challenges) [33] that they may or may not have encountered in their teams. They are asked to specify how big the challenge was to them on a 5-point Likert scale from “not challenging at all” to “extremely challenging” [65]. The challenges are described by a general statement followed by possible examples of how this might have happened; however, the examples are not intended to describe the only way the statements may be true. Students are told, “If the statement is true for you, and the example is not exactly how it happened in your team, please still rank the statement as you experienced it.” At the end of this section, students are asked to choose “what you think was the biggest challenge in your team” from a drop-down list.

In the third section, students are asked to indicate any regulation strategies they used to overcome their greatest challenge. This section is “adaptive” in that the choices of regulations differ depending on the “biggest challenge” the student chose in the previous section. Each of these lists of regulations feature SRL, CoRL, and SSRL options (see Appendix A as one example). Students are prompted to rank on a 5-point Likert-scale the frequency in which they used the various regulations from “never” to “always.” Although there are three distinct theoretical categories of regulations—those where one tries to change oneself (SRL), those where one tries to change others (CoRL), and those where teams work as a group to change together (SSRL)—students are given two categories “What I did...” and “What we did as a team...” to distinguish SRL and CoRL collectively from SSRL.

In the last section of the AIRE survey, students go back to the most important personal goals they identified in the first section and rate the extent to which each of these goals have been achieved. They are also asked to reflect on whether the group played a positive or negative role in the process, and how emotionally satisfied they are with their experience of that particular group learning situation [33].

Together, these sections can be used to determine common themes of teamwork that arise across the entirety of the class.

B. Differences between Fall 2020 and Spring 2021 semesters

We compare Fall 2020 and Spring 2021 semesters to contextualize any demographic differences between the two populations that may influence the interpretation of the results. In addition, providing these data can enable meta-studies that combat normative whiteness and highlight inequities in research [66]. There were 440 students who completed the course at CU in Fall 2020, and 531 in Spring 2021—it is important to note that both of these semesters had significantly lower enrollment compared to the semesters this course was taught in-person prior to the COVID-19 pandemic. Of these students, 407 (92.5% of the class) and 464 (87.4% of the class), respectively, completed a post-course survey that asked for self-reported gender, race and/or ethnicity, major, and year at CU of students in these two semesters. The survey data were collected using another research-based assessment that was administered at the same times as the AIRE survey and memo assignment (Table I). We compared the two student populations using a Mann-Whitney U test and found significant differences in the distributions of class years and gender; specifically, the Spring 2021 semester has significantly more first-year students and men. We are unsure as to why there are more men enrolled in the Spring 2021 semester, but it is typical for first-year students to take the course in the spring as this is “on-track” with the curriculum for engineering and physical science majors.

TABLE I. Self-reported demographic data of 407 students enrolled in the Fall 2020 semester and 464 students enrolled in the Spring 2021 semester of the CU Boulder PHYS-1140 Experimental Physics I course. Categories marked with an asterisk indicate that there are significant differences between the student populations in the two semesters as determined by a Mann-Whitney U test with $p < 0.05$.

Class year*	% of students	
	Fall 2020	Spring 2021
First	8.48	52.16
Second	64.09	28.23
Third	14.46	12.07
Forth	9.48	5.82
Fifth and beyond	3.49	2.16
Gender*	% of students	
	Fall 2020	Spring 2021
Woman	33.14	22.84
Man	61.12	72.84
Other gender or not reported	3.69	4.03
Major	% of students	
	Fall 2020	Spring 2021
Physics and Engineering physics	12.78	16.38
Non-physics engineering	56.76	51.08
Math and other science	22.78	28.88
Other disciplines or not reported	3.68	3.66
Race or ethnicity	% of students	
	Fall 2020	Spring 2021
American Indian/Alaskan Native	0.73	0.22
Asian	9.09	9.05
Black/African American	1.23	0.22
Hispanic/Latino	6.63	3.45
Native Hawaiian/Pacific Islander	0	0.43
White	62.65	67.03
Multirace or ethnicity	9.34	9.27
Other race or ethnicity	3.19	2.80
Not reported	3.81	7.54

In addition, there are some other minor coarse-grained difference between the two semesters. For example, the coding “training” packet was revised in the Spring 2021 semester. However, none of these changes affected aspects of the course intended to address teamwork directly.

C. Analysis methods

1. Qualitative methods

To analyze the students’ memo to future researchers, we used a multilevel coding scheme. We first started with an *a priori* code book containing seven codes: affect, authenticity, coding, community, identity, learning and teamwork, which reflected our motivations for the course. We then isolated the teamwork codes and applied another *a priori*

code book based on the AIRE survey, which had four main codes (i) goals, (ii) challenges, (iii) regulations, and (iv) perceived goal attainment. Each of these codes had a number of subcodes created both *a priori* based on the AIRE survey and emergently during the coding process. A. W. and K. O. separately coded two independent subsets of the responses coded as teamwork (22 teamwork responses) and determined the percent agreement between the two raters to be 93.2%. We report percent agreement instead of Cohen’s kappa because the large number of subcodes, along with the low prevalence of individual codes across the small dataset, can result in unreliable Kappa values [67]. After establishing interrater reliability, the entirety of the dataset was divided and coded by A. W. and K. O. using the code book. All additional emergent codes added after the initial inter-rater reliability were discussed and agreed upon by A. W., K. O. and the rest of the research team.

2. Quantitative methods

When presenting the AIRE survey data, we report on the mean student ranking of the teamwork goals and challenges where the 5-point Likert-scale responses were compressed to a 3-point scale for the analysis. The thickness (height) of the bars represents the standard error of the mean and the range covered by the whiskers represents plus and minus one standard deviation from the mean. In addition, we report on the total number of students who ranked each one of the goals as their “most” and “least” important, as well as the “biggest” challenge. In these cases, the uncertainty is given by the 95% binomial confidence interval. Likewise, when reporting on the perceived goal attainment, satisfaction with teamwork, and the role the team played in helping achieve the student’s goal, we report the total number of students who selected each rank on the 5-point Likert-scale with an uncertainty given by the 95% binomial confidence interval.

D. Limitations

As described above, it is important to note that the qualitative and quantitative data sources come from two different, albeit consecutive, semesters. We conducted coding analysis of over 400 written 1–2 page memos from the students in the Fall 2020 term. The results from the coding analysis inspired the use of the AIRE survey in the Spring 2021 term, which revealed similar findings to the memo coding analysis. Because of the extreme effort involved in qualitatively coding 400 student responses, we were unable to conduct the same intensive coding analysis that was done on the Fall 2020 memo data with the Spring 2021 memos. As we can see from Table I and Sec. III B, there were only minor demographic differences in both the distribution of class years and gender between the two semesters, course revisions, and scheduling changes. This leads to a limitation in our analysis; however, the two semesters were not fundamentally different neither in

terms of course experience nor student population. We are therefore able to compare similarities and differences in our findings from the two data sources and postulate about their meanings to better understand students' teamwork experiences in the CURE. In fact, we believe that the numerous similarities we found between the two semesters in our data speak to the fact students' likely had similar experiences across both semesters.

Another limitation stems from the memo to future researchers having a "predicting" element to the reflections that are not equivalent to past-looking reflections of the AIRE survey. The memo prompt explicitly asked students to make recommendations to future researchers, which led many students to frame regulations that were successful for them and their team as forward-looking recommendations. Because so many students framed their regulations as recommendations, choosing to exclude these from our analysis would have resulted in significant undercounting of student regulations. However, including these responses led to various challenges such as separating co- and socially shared regulations and an inability to directly compare the datasets. We accounted for this as much as possible by explicitly coding only recommendations in the memos that were clearly conducted by the students themselves, rather than merely proposed hypothetically as something different to try in the future. Nevertheless, the regulations *actually* conducted by the students may be overcounted by the coding scheme used for the memos.

In addition, we note that both the memos and AIRE survey responses are student perceptions of their teamwork experiences, rather than observed dynamics of what occurred. This limitation is important as students may be less likely to self-report negative experiences, which might be perceived as a "wrong" answer given the emphasis on teamwork throughout the course [68].

Last, another limitation is that the course was taught during the COVID-19 pandemic and, although all of the students in the course had access to a computer and reliable internet connection, we are unable to account for the multitude of increased stressors that may have impacted them, their ability to participate in the course, or their ability to complete the survey. This may be particularly relevant for this work given the disproportionate impact the COVID-19 pandemic had on increasing isolation and loneliness among college students [69,70].

IV. RESULTS AND DISCUSSION

Despite the numerous teamwork challenges we expected in the PHYS-1140 CURE, in prior work we found that the vast majority of students in the Fall 2020 semester ($> 80\%$) reported that teamwork helped them stay motivated, was fun, helped them learn, and allowed them to conduct the research more successfully [1]. These overwhelmingly positive results, which came from self-reporting on a course survey, motivated us to more deeply explore the team

dynamics that occurred in the course, and the ways in which students were able to overcome the obstacles they faced. We use the lens of SSRL theory, making use of the specific methods and data sources outlined above, to answer four specific research questions about the course.

RQ1. What were some of the goals that students had during teamwork?

RQ2. What were some of the challenges that students faced during teamwork?

RQ3. What regulations did students use to overcome teamwork challenges? In particular, how much did students rely on self-regulated learning, co-regulated learning, and socially shared regulated learning strategies?

RQ4. To what extent did students achieve their goals through teamwork?

In this section, we present and discuss findings to answer these four questions. Teamwork goals are discussed in Sec. IV A (RQ1), challenges in Sec. IV B (RQ2), regulation strategies in Sec. IV C (RQ3), and perceived achievement of goals in Sec. IV D (RQ4). We argue that these more detailed results are broadly consistent with the prior finding that students had largely positive teamwork experiences in the course.

Finally, we provide a discussion on unanswered questions still remaining after our analysis (Sec. IV E) such as, did course design elements lead to the success in teamwork? If so, which were most beneficial? And how did students positive teamwork experiences effect their views on the nature of science, sense of belonging in the scientific community, science identity?

A. Student goals for teamwork

Students enter collaborative learning situations with a variety of both conscious and unconscious personal goals. These goals vary based on a wide range of dimensions, such as specific group activity, explicit goals set by the instructor, mood, and past experiences with group work. According to SRL theory, these goals often act as personal metrics of success for students. In other words, throughout the learning processes, students monitor the achievement of these goals and either regulate their learning such that they can better achieve the goals, or change their criteria for success [33]. However, in group work situations, students must balance their personal goals with the success of the group, ideally by setting goals that embrace, or are at least compatible with, the collaborative environment. This was particularly true in the PHYS-1140 course since team assignments accounted for approximately 50% of the grade [1]. Hence, a good place to start, as we evaluate the success of our own learning goal (that students have positive teamwork experiences in the course), is by analyzing the

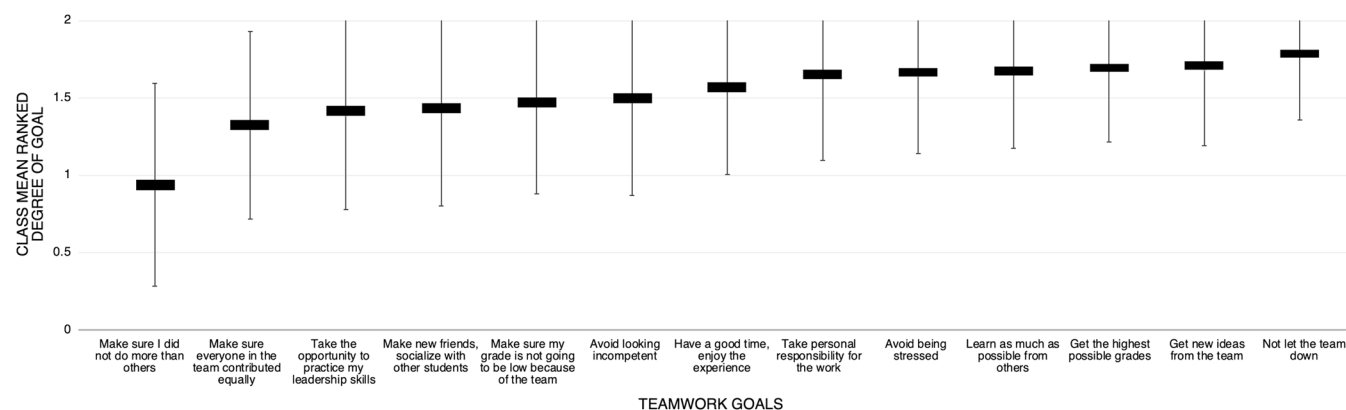


FIG. 1. Students were asked to rank the “degree of importance” of their teamwork goals on the AIRE survey. We plot the mean reported value of all the respondents where 2 is “extremely” or “very” important, 1 is “moderately” or “slightly” important, and 0 is “not important at all.” The whiskers represent plus and minus 1 standard deviation and the height of the bar is the standard error of the mean.

nature of the goals students identified and the extent to which those goals reflected a positive perspective on teamwork. After all, not all goals *concerning* teamwork reflect a positive attitude *towards* teamwork. Goals like “make sure my grade is not going to be low because of the team” or “make sure I do not do more than others” show a clear awareness of the team environment, but suggest that students view it as an obstacle to overcome rather than an opportunity to build upon.

In this section, our primary data sources are the student responses to the AIRE survey (perhaps understandably, students tended not to verbalize their personal goals when writing their memos to future researchers). The AIRE survey provides us with two tools to explore the students’ goals. First, from a list of thirteen possible goals, students were asked to score the importance of each goal on a Likert scale. The class averages of these scores (on a scale of zero

to two; see Fig. 1) give us a sense of the overall level of importance of the goals across the students. Secondly, students were asked to choose which of the thirteen were their most important and least important goals, which gives us a sense of the goal they judged to be most significant relative to the others, regardless of the absolute level of importance (Figure 2).

To begin, we see that students identified a broad range of teamwork goals as being important. On average, student respondents ranked six of the thirteen goals listed in the AIRE survey as extremely or very important to them and almost 80% of the student did not rank any of the listed goals as not important at all. This can be seen in Fig. 1 in which each of the goals has a quite high average importance score.

The goal to “not let the team down” had the highest average importance score ($\bar{x} = 1.79$), indicating a broad

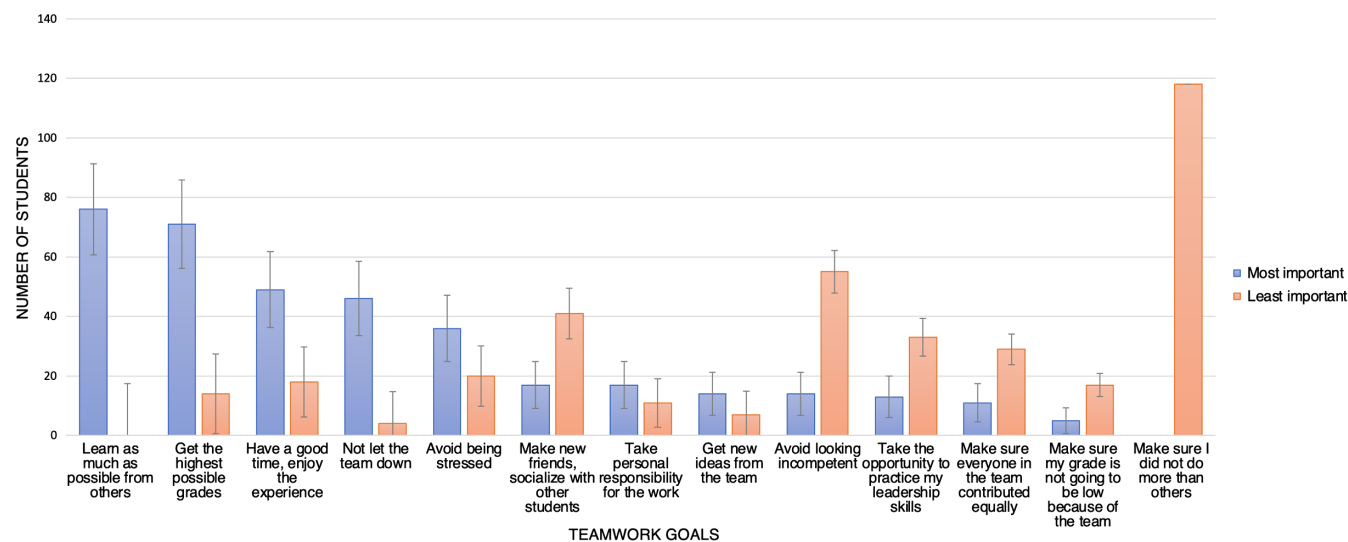


FIG. 2. Most and least important goals as reported by students. Students were asked to choose their “most important” and “least important” goals from the list of goals shown on the x axis. The blue bars represent the number of students reporting that goal as their most important and the orange bars represent the number of students reporting that goal as the least important. Uncertainties are calculated using the binomial confidence interval with $\alpha = 95\%$.

consensus that this very team-oriented goal was important. Interestingly, however, it was not the consensus choice as the *most* important (Fig. 2)—that went to “learn as much as possible from others” (chosen by 20.6%) followed closely by “get the highest grade possible (19.2%). The first of these is clearly consistent with a positive commitment to teamwork; the latter is ambiguous, as it could represent a highly individual goal, or a more group-oriented goal (as in, “my goal is that all of our team members get the highest possible grade”). Either way, the overall picture suggest that, while students balanced an array of teamwork goals, the ones at the forefront of their minds reflected a desire to make their teamwork experience positive and productive.

On the other end, not a single student chose the goal “make sure I did not do more than others” as their most important goal, and 32.0% of students selected this goal as their least important. “Make sure everyone in the team contributed equally,” and “make sure my grade is not going to be low because of the team” also received very low shares of the vote for most important, and were among the lowest for average importance scores. We found these results surprising, as fear that other students will seek to coast through a group project without contributing their fair share, sometimes called “social loafing [71],” is common for students coming into a teamwork experience. We view the relative unimportance of these goals as further evidence that students had positive teamwork experiences in the course. Recall that the AIRE survey was administered at the end of the semester, so that even if students entered with these kinds of concerns, they may have been assuaged by positive teamwork experiences throughout the semester, leading students not to indicate these goals as important in their retrospective assessment.

B. Teamwork challenges

Collaborative learning environments can pose an array of challenges for students [45,72]. Some of these challenges stem from the very nature of collaborative learning as a social environment—below, we outline challenges that are probed in the AIRE survey.

Irreconcilable personal goals: A team can face challenges due to misaligned individual goals, priorities, and expectations [44,45]. For example, goals may be impossible to reconcile if one team member prioritizes minimizing effort and another has the goal of achieving the highest grade possible.

Differing styles of work and communication: Team members may have different styles of working or different ways of interacting and communicating. For example, some people prefer direct communication, while others in the group find the direct communication style confrontational; this can be especially challenging for culturally mixed groups [46].

Differing levels of commitment, concentration, or standards: Different levels of commitment, concentration,

or standards of work among members can create challenges during teamwork [47]. For example, one student may have external commitments that limit their ability to participate fully in the group project.

Differing levels of prior knowledge, understanding of concepts, or power: Cognitive differences (i.e., prior knowledge or ways of thinking) had been shown to be a challenge when working in groups [48]. For example, in a group discussion students may use the same technical terms, but if their understanding of the underlying concepts differ, then what they mean could be quite different. The power structure of the team can also be affected by the perception of differing abilities. For example, some team members who are perceived to have a “high ability” at performing a task compared to the rest of the group may dominate group activity and become communication centers [48].

As with the students’ goals, the AIRE survey gives us two kinds of information about the challenges students reported: a class-average score for the degree of challenge (Fig. 3, on a scale of zero to two), and the class responses when asked which of the challenges they faced was the biggest (Fig. 4). For this section, however, we also draw on our coded analysis of the students’ memos to future researchers, which frequently mentioned challenges faced in the context of giving advice.

In Fig. 4, we see that the most common choice ($17.6 \pm 3.9\%$ of the respondents) for the biggest challenge was “differing in our understanding of the concepts or tasks.” This type of challenge seems both inevitable in a large-enrollment class, and also consistent with the possibility of students having positive teamwork experiences, since it represents a shared challenge that all could work to overcome.

This challenge was frequently brought up in the memos to future researchers from the Fall 2020 term when students discussed concerns about collaboratively coding in a group with significantly different coding expertise. Recall that we tried to form teams with similar self-reported coding confidence, though because of other restrictions and variations in how students self-reported, this did not always result in teams with balanced coding skills, and many referenced these imbalances as a being a challenge, at least initially. One student wrote

I was very apprehensive about coding as I had very little experience and I felt very out of my depth when I compared myself to the rest of the class, but it didn’t take long for me to realise [sic] the many resources at my disposal to help me; the TA, pre-lecture videos, and my team.

For this student, differing in our understanding of the concepts or tasks was initially a challenge for them, which led to apprehensiveness during their group coding; however, they were able to overcome this challenges by using

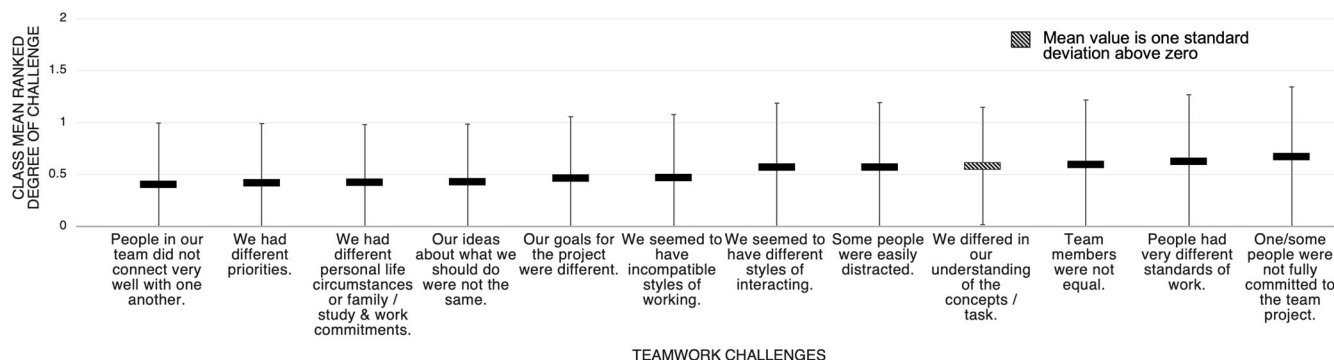


FIG. 3. Students were asked to rank the “degree of difficulty” for each of the challenges on the AIRE survey. We plot the mean reported value of all the respondents where 2 is extremely or very challenging, 1 is moderately or slightly challenging, and 0 is not challenging at all. The whiskers represent plus and minus 1 standard deviation and the height of the bar is the standard error of the mean.

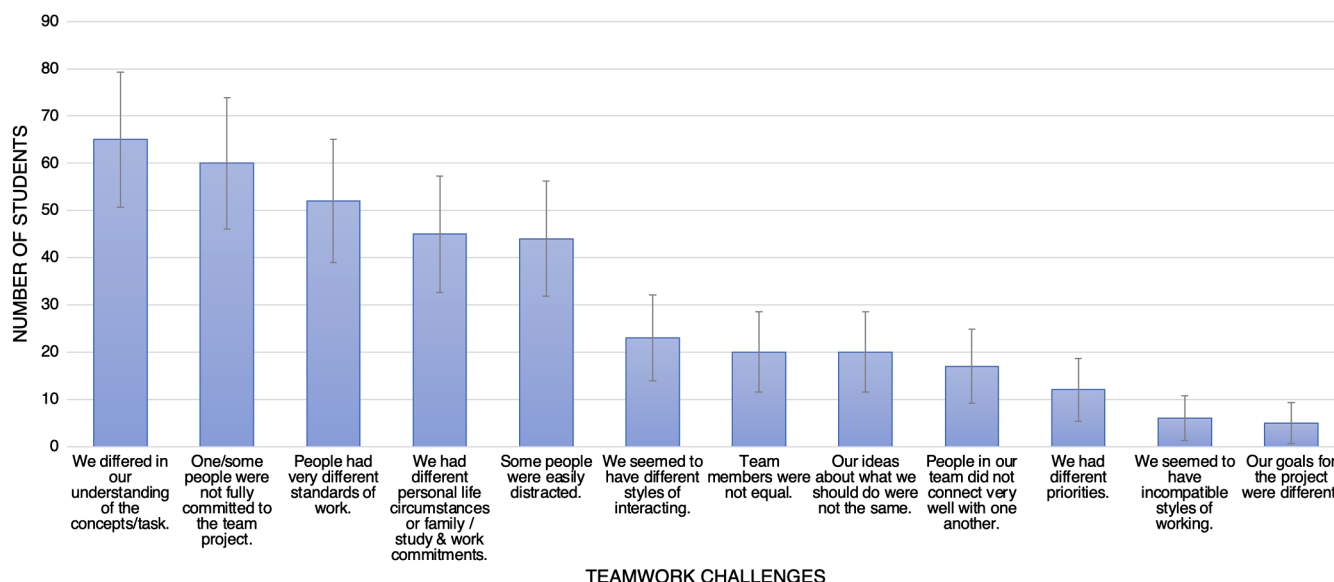


FIG. 4. Biggest challenges as reported by students. Students were asked to choose their “biggest” challenge from the list of goals shown on the x axis. Uncertainties are calculated using the binomial confidence interval with $\alpha = 95\%$.

course resources including the knowledge from other members of their team. In the end, what was initially a challenge turned out to be an advantage for this student.

Interestingly, the second-most popular choice of biggest challenge from the AIRE survey was “one or some people were not fully committed to the team project,” at $16.3 \pm 3.8\%$, followed by “people had very different standards of work,” “we had different personal life circumstances or family or study and work commitments,” and “some people were easily distracted” with $14.1 \pm 3.6\%$, $12.2 \pm 3.3\%$, and $11.9 \pm 3.3\%$, respectively. In other words, a majority of the class, representing approximately 54% of survey participants, chose a “biggest challenge” that fell into the category of “differing levels of commitment, concentration, or standards.”

At first glance, this seems to reflect substantial negativity among the students regarding their teamwork experiences,

and it seems contradictory with the prior finding that very few students expressed teamwork goals geared towards policing social loafing. But this apparent conflict shows the utility of the AIRE survey, and the value of using multiple forms of data when attempting to quantitatively assess teamwork in a physics lab. Clearly, students identified this category of challenges as significant *relative to the others listed in the survey*. But how significant did they consider them to be in more absolute terms? The class average score for the challenge “one or some people were not fully committed to the team project” was 0.67 ± 0.04 , putting it in the middle of the range between not challenging at all (a score of zero) and slightly or moderately challenging (a score of one). In fact, of all 12 challenges presented in the AIRE survey, none received an average score above one (see Fig. 3). Similarly, 76% of respondents did not list *any* of the challenges as being extremely or very challenging.

The picture that emerges is that students did see various forms of social loafing and unequal effort as a challenge, but not a particularly significant one. This is also consistent with what we see in the memos to future researchers. In this context, the advantage of the memos is that they allow students to identify challenges they faced beyond those listed in the AIRE survey. In the memos, we frequently see students discussing collaboration challenges, but of the 217 students who discussed team-related challenges, only 12 students referenced challenges related to differing commitments, concentration, standards or other sentiments related to social loafing. This was further emphasized in the responses to a reflection question asked of students in the Fall 2020 semester (see Table II). In the final week of class, students were asked to report “How many of you were fully prepared for the teamwork most of the time this semester?” and “How many of the team members participated actively most of the time this semester?” Table II shows that, in their reflection responses, the vast majority of students in Fall 2020 felt that their team was fully prepared and participated actively.

In fact, in the memos, most of the teamwork challenges reported by the students were related to the environment or other external factors (e.g., class modality or internet connectivity issues during an online course)—challenges which were not probed by the AIRE survey. These external factors commonly appeared either as general statements about the difficulty of online course work or specific challenges related to collaboratively coding using the Google Colab [73] environment—37.8% students who completed the memos (and 74.2% of the students whose memos brought up specific challenges) reference technological barriers. While these goals directly relate to

teamwork (for example, by placing limitations on how and when students could collaborate on a particular task), they do not by themselves suggest that the students had negative teamwork experiences. Indeed, unlike “social loafing” challenges, which inherently arise from the behaviors of a fellow student, these environmental challenges can potentially represent a “common enemy” external to the group that students can work together to overcome. Consider, for example, the sentiments shared by one student, who reflected that, because the online environment made it difficult to learn, future students should rely on their team more heavily than normal:

Some advice I can offer for you is to work extremely hard as a team. This is the only reason why my team and I survived since during online school, learning can be very tough.

This raises a clear question—was teamwork successful in this class despite the pandemic, or did some of the environmental challenges of the pandemic lead to closer bonding of the teams? A further discussion of this question can be found in Sec. IV E.

C. SRL, CoRL, and SSRL to overcome teamwork challenges

Now that we have identified challenges students faced, we explore how they dealt with these challenges. The AIRE survey features an adaptive section inquiring about students’ use of regulations to respond to challenges. Based on what the students indicated was their biggest challenge, they are presented with 18 possible regulations, which are unique to that challenge. Students are then asked to rank the frequency in which they used each regulation to overcome their challenge from never to always. The 18 regulations presented can be subcategorized into six that feature SRL, six that feature CoRL, and six that feature SSRL (see Appendix A as an example). Although regulations differ between challenges, many are still quite similar. For example, in response to the challenge “we differed in our understanding of the concepts or task” one possible regulation given is “I tried to understand that the others were not simply trying to be difficult but they had different understandings.” For the challenge “One or some people were not fully committed to the team project” the corresponding regulation was written as, “I tried to understand that the others were not simply trying to be difficult but they had different priorities.”

We can therefore use these subcategorizations to show overarching trends in students regulatory practices. For example, Tables III and IV show us that students in the course most often used all three types of regulations; however, while socially shared regulations and self regulations were used by nearly all the students, co-regulations were less common.

TABLE II. Student responses to the final reflection question in the Fall 2020 semester about team participation and preparedness. The uncertainty is given by the 95% binomial confidence interval with $n = 403$.

Number of students in team	How many of you were fully prepared for the teamwork most of the time this semester?
All	$81.1 \pm 3.8\%$
All but one	$12.7 \pm 3.2\%$
Half	$6.0 \pm 2.3\%$
Only one	$0.2 \pm 0.5\%$
None	$0.0 \pm 0.0\%$
Number of students in team	How many of the team members participated actively most of the time this semester?
All	$75.2 \pm 4.2\%$
All but one	$17.9 \pm 3.7\%$
Half	$6.2 \pm 2.4\%$
Only one	$0.7 \pm 0.8\%$
None	$0.0 \pm 0.0\%$

TABLE III. The percentage of students ($n = 369$) who engage in zero, one, or multiple of the regulation types: socially shared, self-, and co- to overcome their teamwork challenges. The error is given by a 95% binomial confidence interval.

Types of regulations used	Percent of the class
0	0.8 ± 0.9
1	2.2 ± 1.5
2	23.0 ± 4.3
3	74.0 ± 4.5

The full dataset provided by this section of the AIRE survey allows us to look in detail at both the number of regulations used by students within a certain subcategorization and the frequency with which they reported using them. The full breakdown of these data are provided in Appendix C, but let us now examine the particular case of the class consensus biggest challenge (which was “we differed in our understanding of concepts or task”) to see how such data can be useful:

Figure 5 shows all of the possible regulations given on the AIRE survey for the respondents who chose this particular challenge as the biggest ($n = 65$). The subfigures are divided into the self-, co-, and socially shared regulations subcategories and we present the mean reported frequency and standard error of the mean for each of the regulations.

Students who self-regulated mainly did this by trying to (i) “become more flexible,” (ii) “understand that the others were not simply trying to be difficult but they had different understandings,” and/or (iii) “accept the situation, realizing that some people had different understandings.” Rarely did students report more concerning kinds regulations such as pulling themselves away from the course and only contributing “the strict minimum” or choosing to interact with only a subset of the team that had the “same understanding as me.” We conclude that students often approached their teamwork challenges with empathy and typically continued to work together as a unit.

Compared to the other categories, co-regulations (which include some potentially negative, confrontational regulations) were only “rarely or sometimes” used by the students to overcome the challenge of differing understandings, Fig. 5(c). Almost none of the students told someone on

their team “that it would be better if they did not contribute much” or that they needed “to change their understandings or face the consequences.” Students did sometimes regulate this challenge by telling the others that they needed to “accept that some people had different understandings of the concepts and the tasks,” “be more open in order to find a compromise or solution for this situation,” or “be aware of our different understandings.” Students sometimes also “tried to convince someone that the others were not simply trying to be difficult and that we could sort out the situation.” These co-regulations can be viewed as potentially mediating between two other teammates who were more confrontational in regards to the challenge.

Finally, for the socially shared regulations, most students said that they used the regulation: “As a team we decided to sort out the situation together and found a way to complete our work in the best possible conditions” to overcome the challenge. But many other socially shared regulations were employed as well; as we can see in Fig. 5(a) the students said that they often engaged in five of the six socially shared regulations. We also note that the least-frequently used socially shared regulation was “we sorted out the situation by agreeing that some people would not contribute much.” Once again, this potentially speaks to the positive teamwork environment that was present in the course, and this may be also true of the Fall 2020 class whose students’ commonly expressed the goal of wanting to learn as much as possible from others.

We can also analyze the students choices of regulation based on their comments in the memos to future researchers, by coding which types of regulations were mentioned. In the memos, we were unable to separate co- and socially shared regulations (see Sec. III D) but we found that, of the 527 regulations referenced by the students, 95% were CoRL/SSRL and only 5% used SRL. Furthermore, when students discussed regulations using SRL, they often also included a regulation using SSRL/CoRL. For example,

Speaking of team mates, it’s very important to harmonize as a group because that is how you get the most out of your research. If you are going to be absent or unable to work in the lab be sure to make it up in later sections. Your group needs to rely on each other in order to succeed. I’d also recommend delegating tasks for each member each week of the lab so you can efficiently gather all the needed information and understand it as well.

This student discussed the importance of harmonizing as a group in order to succeed in getting “the most out of your research.” They suggest that a future researcher should do this by using the self-regulation of making up the work in a later week if they are absent—implying that one may not be able to participate equally every week of class due to external factors, but it is important to still try to keep the total work contributed to the team as equal as possible. In addition, they suggest that a socially shared or co-regulation is needed to achieve this goal—“delegating tasks for each member each week of the lab.”

TABLE IV. The percentage of students ($n = 369$) who reported that they used socially shared, self-, and/or co-regulation on the AIRE survey to overcome their teamwork challenges. The error is given by a 95% binomial confidence interval.

Type of regulations used	Percent of the class
Self	95.9 ± 2.0
Co	74.8 ± 4.4
Socially shared	97.8 ± 1.5

TABLE V. Regulation types referenced in the memos to future researchers. The percent of students is calculated from the total number of memos, $n = 403$.

Regulation	Number of student references	% of student references
Divided the work	154	38.0
Communicated	82	20.3
Used zoom features	71	17.6
Worked together	53	13.2
Got to know each other	35	8.7
Taught and learned from each other	30	7.4
Discussed goals or priorities	20	5.0
Asked questions	19	4.7
Involved everyone	9	2.2
Compromised	2	0.5
Other	26	6.5

Of the regulations discussed in the memos, many used other techniques that were not specifically referenced in the AIRE survey (see Table V). The most commonly referenced regulation was “dividing the work,” which allowed them to achieve many of their teamwork goals, such as equity of work distribution, getting the work done, and learning from others. More examples of these can be found in Appendix B, Table VII. Unsurprisingly, given how frequently they cited external challenges, students also referenced a number of regulations aimed at redressing these external obstacles in their memos.

On the whole, we see that students engaged in a wide variety of regulations to address the teamwork challenges they faced, but generally avoided the kinds of regulations one might associate with a more dysfunctional and confrontational team environment. Socially shared regulations, which mutually and collectively involve the whole group, were widely cited in both the AIRE survey and the memos. Self-regulations were also commonly identified on the survey, but were interestingly and notably absent from the memos. One possible interpretation of this would be that, while students in practice found themselves employing large amounts of both SLR and SSLR (as reported on AIRE), they more commonly identified the SSLR regulations as the better “advice” to recommend to others. While this is speculative, it would again suggest a mindset focused on collaborative effort as a key to success in the class.

We view the students’ broad use of constructive regulations and primary focus on socially shared regulations as a success relative to the stated goals of the course. We can consider a few different possible reasons for these outcomes: (i) students may have come in to the course with prior knowledge about best ways to navigate teamwork, (ii) the challenges they experienced were minor and easy to resolve without confrontation, or (iii) educational interventions in the course such as the group formation, regular metacognitive

reflections, and the teamwork training contributed to students choices. While the first seems to be the least probable—since the majority of the students were freshmen and sophomores and likely had little prior experience working collaboratively on research projects at the collegiate level—it is likely that some combination of all three causes is present.

D. Perceived goal attainment

Finally, now that we have identified student goals, challenges, and regulations for these challenges, we can explore if these goals were actually attained in the course. Although we cannot pinpoint the exact cause, we do know that by using regulations, students overwhelmingly achieved their teamwork goals and reported that their group played a positive role in their success. In the final section of the AIRE survey, students were asked to

1. Rate the extent to which they thought their most important teamwork goal was achieved.
2. Rate the extent to which the team played a positive, neutral or negative role helping achieve their most important teamwork goal.
3. And rate how personally satisfied they were with their teamwork experience this semester.

We see in Fig. 6 that $85.6 \pm 9.9\%$ of the respondents felt that their teamwork goal was fully or mostly achieved. Likewise, $82.9 \pm 9.3\%$ and $84.0 \pm 9.9\%$ of the respondents, respectively, felt that their team played a positive role in achieving their teamwork goal (Fig. 7) and that they were satisfied with their teamwork experience (Fig. 8). We saw similar findings in the memos to future researchers during the Fall 2020 semester. In the memos, students’ expressed that teamwork lead to the achievement of many goals such as conducting successful research (58 students), getting work done (81 students), learning skills and concepts (102 students), and having fun (58 students). Appendix B, Table VIII highlights a few quotes that speak to the goals students felt that they achieved through their teamwork in the class.

E. Implications and remaining questions

The results presented above for all four of our research questions are consistent with our preliminary finding that students generally had positive teamwork experiences in the course [1]. Moreover, they help us to understand some of the specific ways that this came to be, and how students were able to navigate the inevitable challenges that arise in a group-work setting. We saw students setting a broad array of teamwork goals, indicating significant intentional engagement with teamwork as a concept. Along these lines, more negative goals like “making sure not to do more than others” were considered fairly unimportant. Students did identify various teamwork challenges they faced, but their average scores on the AIRE survey put them all in the range between “no challenge at all” and “slightly or moderately challenging.” In fact, rather than identify the

kinds of challenges that would indicate friction or competition between teammates, the most commonly identified problems were logistical and externally imposed. In response to these challenges, students reported using a broad array of regulations to address them, and primarily reported (and recommended to future students) SSRL, which in many ways was the most desirable from the perspective of our teamwork goal. Finally, a large majority of the students reported that their own teamwork goals were achieved, and that their teams played a positive role in this success.

The comprehensive picture painted by these results is extremely positive for our overall teamwork goal. However, there are still quite a few questions which remain unanswered, such as,

1. Did course design elements lead to the success in teamwork? If so, which were most beneficial?
2. Did the research itself (i.e., being a CURE) change how students viewed teamwork and its importance?
3. Were some of the positive feelings toward teamwork in this course because of contrasting isolation that students may have experienced in other classes during pandemic or despite the pandemic?
4. How did students' positive teamwork experiences affect their views on the nature of science, sense of belonging in the scientific community, and science identity?

Although we cannot answer all of these questions, we propose some potential explanations, which we hope can also provide ideas or possible recommendations for the designers of future lab courses and CURES.

To begin with, we observe that past studies have noted the importance of the first day of class, as it can be used to mitigate student concerns, establish course norms, set expectations, communicate the importance of course activities, and increase student motivation [74–79]. Engaging students in teamwork training on the first day of class, described in Sec. II C, may have not only provided students with useful tools to overcome common teamwork challenges, but also set the tone for the course and prepared students for the importance and intensity of teamwork throughout the semester. In addition, many of the of the weekly reflection questions in the course asked students about their teamwork experience (4 out of 12), potentially encouraging the regulatory process—reflecting on teamwork challenges and thinking of solutions outside of the lab time. Lastly, the teams were created specifically to avoid groups in which only one student did not identify as a man and to form teams whose members reported a similar level of coding confidence. These two choices were made to address three potential teamwork challenges in this course: (i) lower performance and poorer social cohesion of women in male-dominated groups [57], (ii) gendered division of roles seen more commonly in unstructured physics labs like ours [58], and (iii) unequal distribution of work due to different prior coding knowledge. Although differing understandings was reported on the AIRE survey as one of the largest teamwork challenges faced by students, it

may have been a much more significant challenge if teams were not formed in such a way. Ultimately, however, we do not know which of these choices impacted students most significantly, or indeed whether they contributed more than external factors such as the students' prior knowledge about successfully navigating teamwork scenarios.

The teamwork aspects of the course may also have been successful due to the very nature of the CURE environment itself. During the first lecture (a 15–20 min online, asynchronous video which students were required to watch prior to lab) the instructor emphasized the importance of collaboration as a scientific practice giving examples of a large collaboration like CERN, which involves thousands of people to the 10-person, primarily graduate student lab that the instructor runs at CU Boulder. The instructor further stated that one of the “most sought-after skills” in science industry jobs is the “ability to work on an interdisciplinary team.” Working toward a unified research goal in an authentic way may have allowed students to bond and find motivation that traditional labs may not offer. In addition, the PHYS-1140 CURE featured a variety of authentic practices such as peer review and two group meetings with the principal investigator, which may have emphasized the importance of collaboration as a part of physics.

Finally, we consider, as a possible explanation, that the overwhelmingly positive teamwork results may have actually been because of the pandemic environment rather than despite of it. Certainly, the pandemic factors posed added difficulties for the students. Recent work by Wildman *et al.* (2021) [34], which analyzed open-ended survey responses of students working in project teams during the pandemic, saw that there were

1. Increased internal distractions, forgetfulness, and procrastination;
2. Individual challenges, exacerbated by the pandemic, that had impacts on the larger team;
3. New challenges, such as navigating geographical differences and difficulties communicating; and
4. Increased progress disruptions, ambiguity, and loss of morale within the team.

But despite these negative findings, they also found some positive changes in teamwork dynamics the COVID-19 pandemic, such as more efficient meetings or increased communication and empathy [34]. In addition, at the end of the Fall 2020 semester, students were asked to compare the PHYS-1140 course to other courses they were taking concurrently. One question asked for students to rank their agreement with the statement, “as compared to other classes you took this semester, this class provided me with a community” to which $66.1 \pm 42\%$ of the class agreed or strongly agreed. The strong emphasis on teamwork and collaboration in this course may have felt like a reprieve compared to the isolation felt in other online courses during the COVID-19 pandemic [69,70], particularly for the first- and second-year students that represented the majority of the class. Overall, the desire to take advantage of the

teamwork environment presented in our CURE may have encouraged students to embrace cooperation and collaboration to a greater extent than would have been possible under more normal conditions.

One major remaining question is how the positive teamwork experiences in this course may have affected students' views on the nature of science, their sense of belonging in the scientific community, and their science identity. A major work by Auchincloss *et al.*, which defined the key aspects of CUREs, described collaboration as one of the defining features, stating that, "science research increasingly involves teams of scientists who contribute diverse skills to tackling large and complex problems [3]" and that "group work is not only a common practical necessity, but also an important pedagogical element of CUREs because it exposes students to the benefits of bringing together many minds and hands to tackle a problem [3]." It is not only important for students to develop valuable teamwork skills, but also to understand the social nature of scientific work. The social aspect of science has been identified as a key element of the nature of science by Osborne, Collins, Ratcliffe, and Millar [80]. They write that when teaching the theme of cooperation and collaboration in science, it is important to "stress the social processes in science, as this was an aspect too often overlooked in school science" especially since science is often viewed as "the retreat of the lone genius [80]." In a postcourse survey administered in Fall 2020 of the PHYS-1140 CURE, $84.4 \pm 3.2\%$ of students agreed or strongly agreed that "after completing this course, I better understand the process of conducting scientific research" and $88.1 \pm 2.7\%$ agreed or strongly agreed that "after completing this courses, I believe research is inherently collaborative [1]." However, more work is needed to identify how the teamwork aspects of this course may have effected these results. Furthermore, we believe more research needs to be done to better understand how teamwork—specifically in CUREs—impacts students' views and beliefs on the nature of science, their sense of belonging in the scientific community, and their science identity.

V. CONCLUSIONS

After the Fall 2020 semester, our preliminary analysis [1] found that students gave overwhelmingly positive feedback about teamwork experiences with over 80% of the class reporting that teamwork helped them stay motivated in the course, learn, and more successfully conduct their research. Through this work, our goal was to further explore how students engaged in teamwork during the PHYS-1140 CURE to better understand their positive feelings. We did this by using data from both the AIRE survey [33] and the written responses to the memos for future researchers, a final assignment of the course, to understand students' goals for teamwork, whether these goals were achieved, what teamwork challenges students needed to overcome, and how they overcame those challenges.

A. Summary of findings

We found that in the Spring 2021 term the students overwhelmingly felt that they had achieved their teamwork goals (85.6%) and felt that their teams played a positive role in their success (82.9%). Two possible reasons for this are that (i) not many students reported that they faced major teamwork challenges with 76% of the respondents on the AIRE survey not listing any of the challenges as extremely or very challenging and (ii) the challenges that students did face were mostly addressed through productive, socially shared regulations such as deciding to "sort out the situation together" and finding a way to "complete our work in the best possible condition." We found similar results in our analysis of the memos to future researchers from the Fall 2020 semester, where students rarely discussed any of the social teamwork challenges that were probed by the AIRE survey, and discussed primarily external factors, such as collaboratively coding in Google Colab [81], as a major challenge. However, despite the external and environmental teamwork challenges, students recommended a variety of productive regulations for the "future researchers" to try that worked successfully for them. Most of these solutions involved socially shared regulations where students worked as a team to divide the work and rotate roles.

B. Future implications

Our course was the first implementation of a CURE in a large-scale introductory physics course. Our primary goals for the CURE were to teach research skills, foster productive teamwork, and give students a positive experience with experimental research. This work can serve as a building block for the development and implementation of more physics CUREs, especially at the introductory level. But in addition, we encourage physics lab instructors to emphasize and teach teamwork as an explicit learning goal in their courses.

Even though collaboration and teamwork are an integral parts of modern scientific practice and seen as a vital skill by the scientific and physics communities, many physics lab courses only passively teach teamwork skills by assigning lab groups. Through this work, we have described some educational interventions—framing in the lectures, syllabus, and canvas website; teamwork training; metacognitive reflection questions; purposeful team assignments; and grading practices—which may help students develop teamwork skills. Likewise, engaging students in authentic collaborative scientific practices such as peer review and group meetings could further emphasize the importance of teamwork within the physics community.

Finally, comparatively few instructors assess teamwork outcomes in their courses the way it has become more common to use assignments and evaluative tools to assess content-based learning goals for the purpose of course improvement. Part of this may be due to the complex, socially driven nature of teamwork itself, which can be

difficult to capture. This paper provides an in-depth example of how one can both qualitatively and quantitatively assess the success of a teamwork goal through the use of multiple assessments and artifacts. This type of data analysis could be used, not only in a CURE, but in any lab course, or perhaps even in a traditional undergraduate research experiences. However, there is a clear need for better quantitative methods for evaluating “successful” teamwork, student learning of teamwork skills, and the impact of teamwork on students understanding of the nature of science and identity within the context of scientific environment. While the AIRE questionnaire used in this work provides valuable information about common teamwork goals, challenges, and regulations, it is not specific to teamwork and collaboration as part of research in physics and the sciences. Furthermore, quantitative methods of evaluating teamwork in physics would be particularly important in large-enrollment classes, such as PHYS-1140, where it may be difficult for instructors to assess teamwork qualitatively using written responses, interviews, or observational data.

ACKNOWLEDGMENTS

We thank the students in the course who took the time to share their experiences with us. This work is supported

by STROBE National Science Foundation Science Technology Center, Grant No. DMR-1548924.

APPENDIX A: EXAMPLE REGULATIONS FROM AIRE QUESTIONNAIRE

We provide an example of regulations to the teamwork challenge “we differed in our understanding of the concepts/task” from the AIRE survey (Table VI). Each regulation is categorized into SRL, CoRL and SSRL—six in each category. All the challenges have similar regulations to the example below, but the language is slightly modified to the specific challenge. Additionally, all the regulations for each of the challenges can be categorized into SRL, CoRL and SSRL with six regulations in each category.

APPENDIX B: EXAMPLE QUOTES FROM THE MEMOS TO FUTURE RESEARCHERS

Example quotes from the student memos to future researchers from the Fall 2020 semester that highlight teamwork goals that students expressed they could achieve through dividing the work (Table VII) and various teamwork goals that students expressed they achieved (Table VIII).

TABLE VI. Example of regulations to the teamwork challenge “we differed in our understanding of the concepts or task” from the AIRE survey. Each regulation is categorized into SRL, CoRL, and SSRL—six in each category.

Type of regulation	Regulations to “We differed in our understanding of the concepts or task.”
Self	I convinced myself that it could actually be a good thing, because ... I tried to become more flexible. I tried to accept the situation, realizing that some people had different understandings. I interacted only with the team members who had the same understanding as me. I tried to understand that the others were not simply trying to be difficult but they had different understandings. I decided to contribute only the strict minimum.
Co	I told the others that we needed to accept that some people had different understandings of the concepts and the task. I tried to explain to others that we needed to be aware of our different understandings. I told someone that it would be better if they did not contribute much. I tried to convince someone that the others were not simply trying to be difficult and that we could sort out the situation. I told someone to change their understandings or face the consequences. I told the others we needed to be more open in order to find a compromise or solution for this situation.
Socially shared	We realized we had to reconcile our various understandings of the concepts or tasks and bring them closer to one another. We tried to adopt the majority’s understanding. We resolved the situation by compromising to accommodate everyone’s understandings. We decided to sort out the situation together and found a way to complete our work in the best possible conditions. We sorted out the situation by agreeing that some people would not contribute much. We accepted that different understandings and we had to organize our work accordingly.

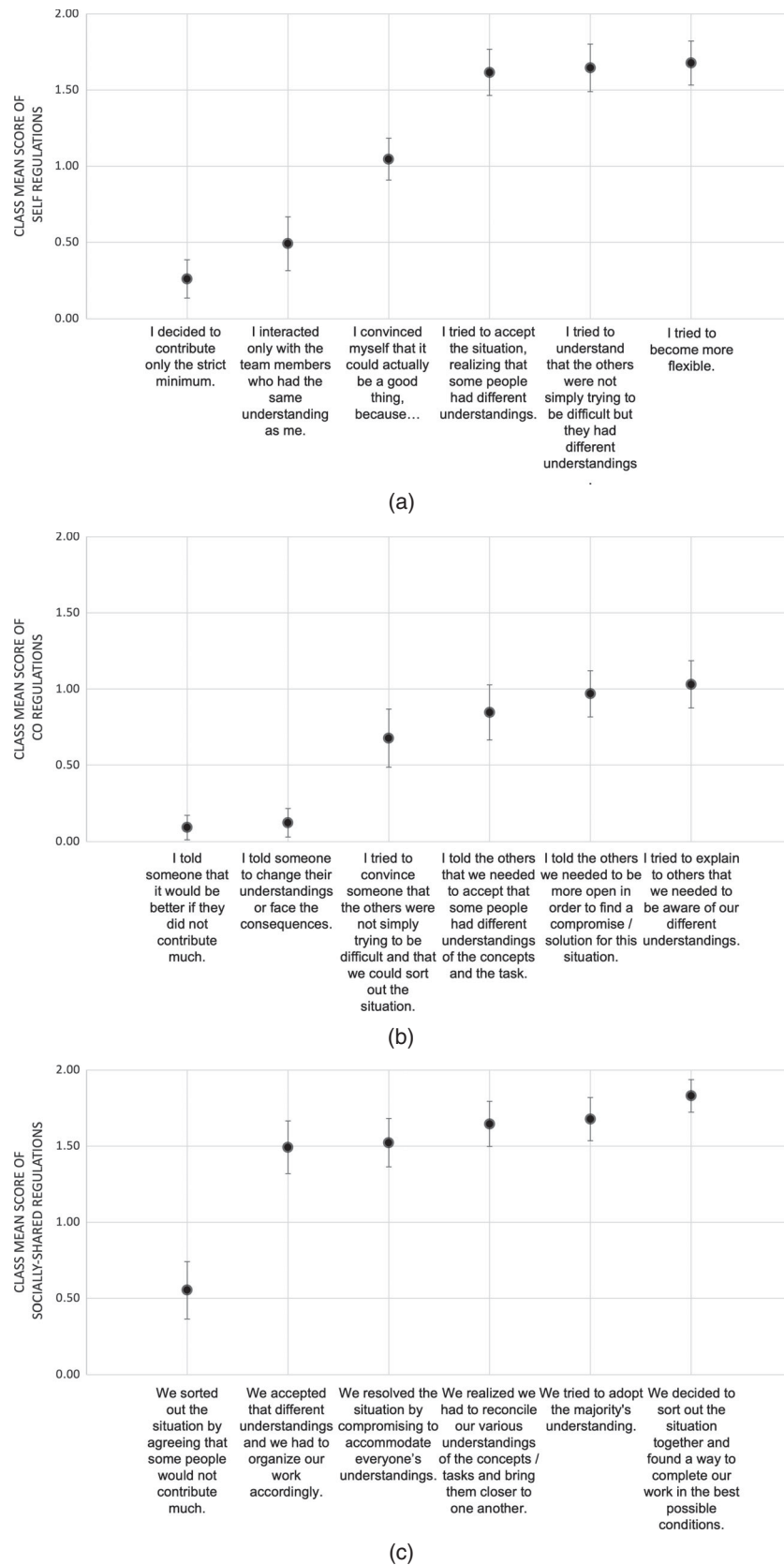


FIG. 5. Students were asked to rank the frequency that they used each of the (a) self-regulations, (b) co-regulations, or (c) socially shared regulations to overcome the challenge of having differing understandings of concepts where 2 is always or often, 1 is sometimes or rarely, and 0 is never. The uncertainty is given by the standard error of the mean.

Rate the extent to which you think your most important goal was achieved:

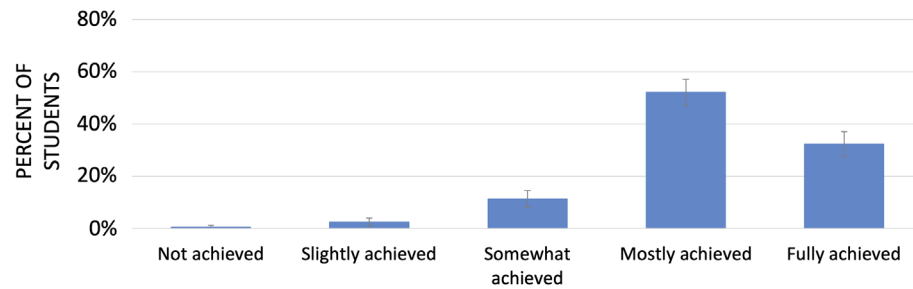


FIG. 6. Student responses ranking achievement of their previously indicated most important goal this semester. Uncertainties are calculated using the binomial confidence interval with $\alpha = 95\%$.

Please rate the extent to which the team played a positive, neutral or negative role in helping you achieve the three things that were important to you:

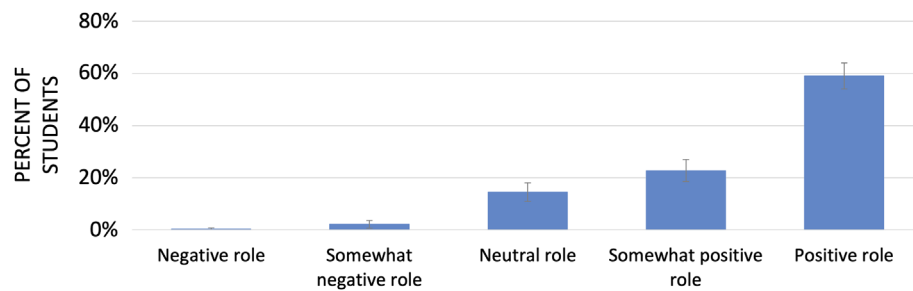


FIG. 7. Student responses ranking the role that the team played in achieving their indicated most important goal. Uncertainties are calculated using the binomial confidence interval with $\alpha = 95\%$.

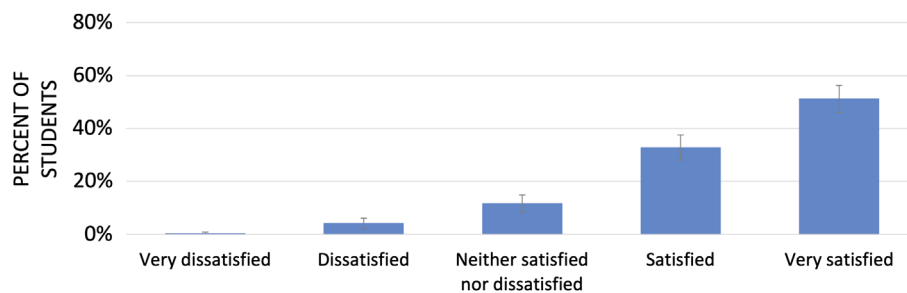


FIG. 8. Student responses ranking satisfaction in their teamwork experience this semester. Uncertainties are calculated using the binomial confidence interval with $\alpha = 95\%$.

APPENDIX C: BREAKDOWN OF REGULATIONS USED IN RESPONSE TO TEAMWORK CHALLENGES

Using the data from section three of the AIRE survey, we can look at the frequency and number of regulations that

students used within a certain sub-categorization (Fig. 9). We see in Fig. 9 that students frequently used all six socially shared regulations and four of the self-regulations presented on the AIRE survey, with varying degrees of frequency. By comparison, when students did engage in co-regulations, they did so often, but used only a few of the co-regulations.

TABLE VII. Example quotes highlighting various teamwork goals that students expressed they could achieve through dividing the work.

Teamwork goal achieved	Example quotes from memos to future researchers
Equity of work distribution	(1) A piece of advice I would give for working on the collabs [sic] would be, have one person write the code in the collab [sic] and share their screen so everyone can give tips or add comments, and have another person write the drop down cells in a separate google doc. Once you are done, have the code writer copy and paste the drop down cell answers into the collab [sic] and then submit it that way. Rotate who does the coding and drop down cells answers every week too so everyone gets a chance to code and answer the questions. (2) We took turns in each role so everyone got a chance to take part in all aspects of the research.
Getting the work done	(1) With regards to teamwork, try to divide up the work as much as possible. Try to make sure no one is left out, and that no one is doing too much. Having good teamwork means you do not have to do as much outside of lab. (2) Teamwork is an important part of the research process and will make the assignments exponentially easier to complete. As long as you stay focused and divvy up tasks in the recitation sessions, you shouldn't have to worry about completing any work outside of class.
Learning	(1) Your group needs to rely on each other in order to succeed. I'd also recommend delegating tasks for each member each week of the lab so you can efficiently gather all the needed information and understand it as well. (2) I would highly recommend making sure everyone in your group is following along with the coding and that only one person work on Google Colab at a time and rotate who is typing every week. This way at least everyone feels like they can participate at some point and everyone has to learn some of the coding because the group will have to collaborate to solve the coding issues.

TABLE VIII. Example quotes highlighting various teamwork goals that students expressed they achieved in the memos to future researchers.

Teamwork goal achieved	Example quotes from memo to future researchers
Conducting successful research	1. The research goal of your group's assignment is to accurately determine peak flare energy and irradiance of a specific solar flare. You will achieve this goal through coding, science, and most importantly, teamwork. 2. Also, the experience of participating in real research and collaborating as a team was valuable to me as an aspiring scientist, because nearly all research is done with a team and requires collaborative efforts.
Getting work done	1. My team and I were able to communicate very well for the entire duration of the semester and we never once had to do extra work outside of the lab period. 2. Even if your team doesn't turn out nearly as good as mine, teamwork and communication are going to make this project so much easier, especially in dividing out the work and completing it most efficiently
Learning skills and concepts	1. However, I had terrific teammates who worked together with me to learn how to effectively use Python and it made the process easier than I could have imagined. 2. When others on your team get stuck or are confused about something you can help them out which will further improve your own understanding of the concepts and formulas.
Having fun	1. My teammates were the best people I could have possibly asked for, and each class period was fun because of the work we were doing together. 2. This semester was enjoyable for me because of my team dynamics. We were focused and did the work we had to do, but we also just talked and had fun as a group.

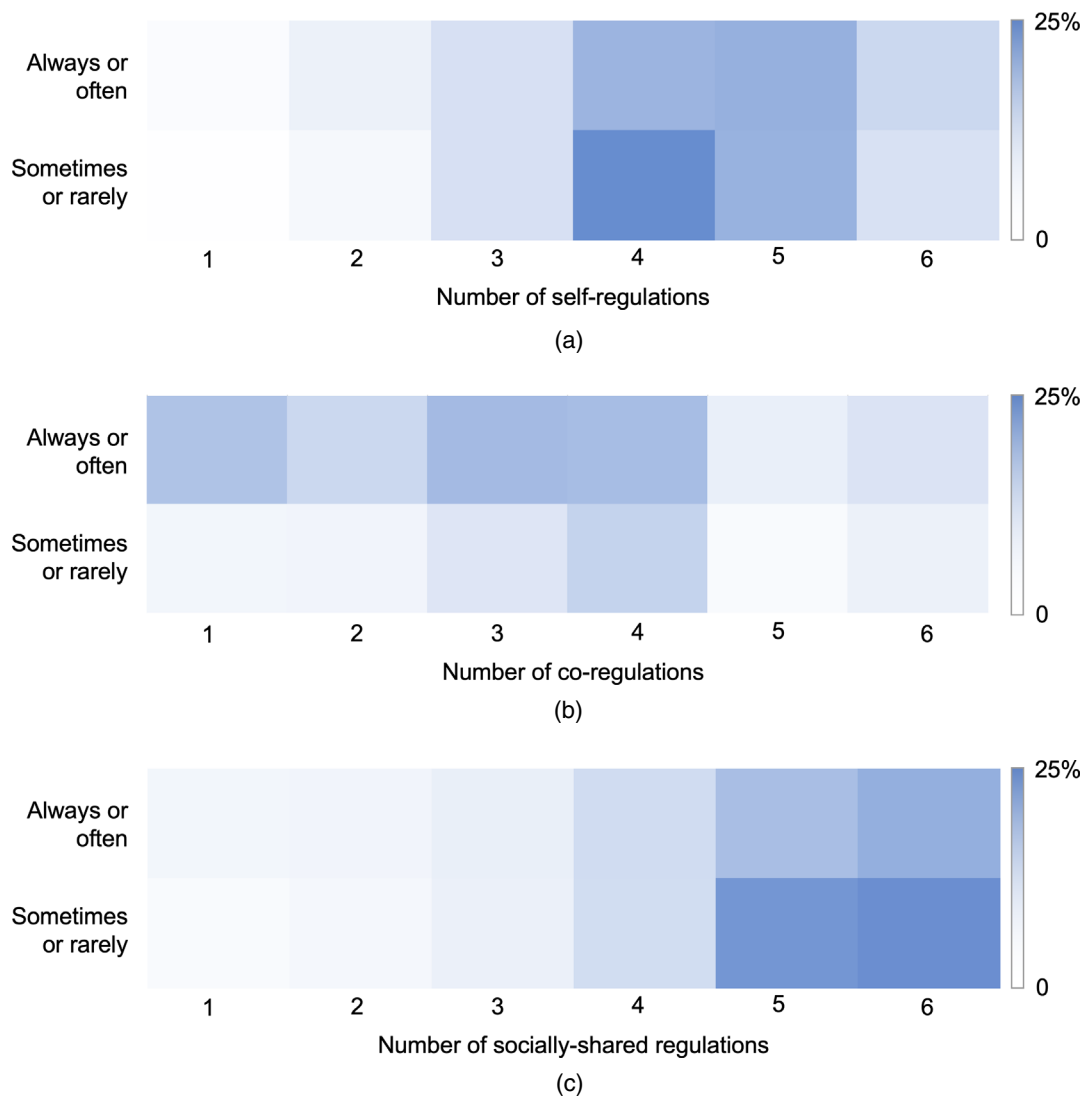


FIG. 9. Students were asked to rank the frequency with which they used each of the 18 regulations to overcome their “largest challenge.” We separate these regulations into subcategories of (a) self-regulations, (b) co-regulations, (c) socially shared regulations. Here, we plot the percentage of respondents who reported using each particular number of regulations for each particular frequency. For example, about 25% of respondents reported that they used four of the self-regulations, but used them only “sometimes or rarely” with $n = 369$). The color of the block gives the percent of respondents who fall into each particular category.

- [1] Alexandra Werth, Colin G. West, and H. J. Lewandowski, Impacts on student learning, confidence, and affect and in a remote, large-enrollment, course-based undergraduate research experience in physics, *Phys. Rev. Phys. Educ. Res.* **18**, 010129 (2022).
- [2] *Undergraduate Research Experiences for STEM Students: Successes, Challenges, and Opportunities*, edited by James

- Gentile, Kerry Brenner, and Amy Stephens (The National Academies Press, Washington, DC, 2017).
- [3] Lisa Corwin Auchincloss, Sandra L. Laursen, Janet L. Branchaw, Kevin Eagan, Mark Graham, David I. Hanauer, Gwendolyn Lawrie, Colleen M. McLinn, Nancy Pelaez, Susan Rowland, Marcy Towns, Nancy M. Trautmann, Pratibha Varma-Nelson, Timothy J. Weston, and Erin

- L. Dolan, Assessment of course-based undergraduate research experiences: A meeting report, *CBE Life Sci. Educ.* **13**, 29 (2014).
- [4] J. F. Kremer and R. G. Bringle, The effects of an intensive research experience on the careers of talented undergraduates, *J. Res. Dev. Educ.* **24**, 1 (1990), <https://psycnet.apa.org/record/1991-13818-001>.
- [5] Heather Thiry, Sandra L. Laursen, and Anne-Barrie Hunter, What experiences help students become scientists? A comparative study of research and other sources of personal and professional gains for STEM undergraduates, *J. Higher Educ.* **82**, 357 (2011).
- [6] David Lopatto, Survey of Undergraduate Research Experiences (SURE): First findings, *Cell Biol. Educ.* **3**, 270 (2004).
- [7] David Lopatto, Undergraduate research experiences support science career decisions and active learning, *CBE Life Sci. Educ.* **6**, 297 (2007).
- [8] Anne-Barrie Hunter, Sandra L. Laursen, and Elaine Seymour, Becoming a scientist: The role of undergraduate research in students' cognitive, personal, and professional development, *Sci. Educ.* **91**, 36 (2007).
- [9] Lisa Corwin Auchincloss, Assessment of course-based undergraduate research experiences: A meeting report, *CBE Life Sci. Educ.* **13**, 29 (2014).
- [10] Marcia B. Baxter Magolda, Identity and learning: Student affairs' role in transforming higher education, *J. Coll. Student Dev.* **44**, 231 (2003).
- [11] David I. Hanauer, Mark J. Graham, and Graham F. Hatfull, A measure of college student persistence in the sciences (pits), *CBE Life Sci. Educ.* **15**, ar54 (2016).
- [12] William B. Wood, Inquiry-based undergraduate teaching in the life sciences at large research universities: A perspective on the Boyer commission report, *Cell Biol. Educ.* **2**, 112 (2003).
- [13] Sophie Pierszalowski, Rican Vue, and Jana Bouwma-Gearhart, Overcoming barriers in access to high quality education after matriculation: Promoting strategies and tactics for engagement of underrepresented groups in undergraduate research via institutional diversity action plans, *J. STEM Educ.* **19** (2018), <https://www.learntechlib.org/p/182980/>.
- [14] Sophie Pierszalowski, Jana Bouwma-Gearhart, and Lindsay Marlow, A systematic review of barriers to accessing undergraduate research for STEM students: Problematising under-researched factors for students of color, *Social Sci.* **10**, 328 (2021).
- [15] Gita Banger and Sara E Brownell, Course-based undergraduate research experiences can make scientific research more inclusive, *CBE Life Sci. Educ.* **13**, 602 (2014).
- [16] Sara E. Brownell, Daria S. Hekmat-Scafe, Veena Singla, Patricia Chandler Seawell, Jamie F. Conklin Imam, Sarah L. Eddy, Tim Stearns, and Martha S. Cyert, A high-enrollment course-based undergraduate research experience improves student conceptions of scientific thinking and ability to interpret data, *CBE Life Sci. Educ.* **14**, ar21 (2015).
- [17] Ramesh Laungani, Colby Tanner, Tessa Durham Brooks, Barbara Clement, Melissa Clouse, Erin Doyle, Scott Dworak, Brad Elder, Kate Marley, and Brett Schofield, Finding some good in an invasive species: Introduction and assessment of a novel cure to improve experimental design in undergraduate biology classrooms †, *J. Microbiol. Biol. Educ.* **19**, 05 (2018).
- [18] James A Klimchuk, On solving the coronal heating problem, *Sol. Phys.* **234**, 41 (2006).
- [19] Benjamin M. Zwickl, Takako Hirokawa, Noah D. Finkelstein, and Heather J. Lewandowski, Epistemology and expectations survey about experimental physics: Development and initial results, *Phys. Rev. ST Phys. Educ. Res.* **10**, 010120 (2014).
- [20] Bethany R. Wilcox and Heather J. Lewandowski, Students' views about the nature of experimental physics, *Phys. Rev. Phys. Educ. Res.* **13**, 020110 (2017).
- [21] Paul W. Gresser, A study of social interaction and teamwork in reformed physics laboratories., Ph.D. thesis, University of Maryland, College Park, 2006, <https://drum.lib.umd.edu/handle/1903/3362>.
- [22] Julian Goñi, Catalina Cortázar, Danilo Alvares, Uranía Donoso, and Constanza Miranda, Is teamwork different online versus face-to-face? A case in engineering education, *Sustainability* **12**, 10444 (2020).
- [23] Ferdinand Keil, David Riehl, and Klaus Hofmann, Converting an undergrad-lab to an interactive e-learning experience that enables student teamwork, in *Proceedings of the 2021 IEEE Frontiers in Education Conference (FIE)* (IEEE, Bellingham, WA, 2021), pp. 1–5.
- [24] Joseph Kozminski, H. J. Lewandowski, Nancy Beverly, Steve Lindaas, Duane Deardorff, Ann Reagan, Richard Dietz, Randy Tagg, Melissa Eblen-Zayas, Jeremiah Williams, Robert Hobbs, and Benjamin Zwickl, AAPT Recommendations for the Undergraduate Physics Laboratory Curriculum Subcommittee Membership., Technical Report (American Association of Physics Teachers Committee on Laboratories, College Park, MD, 2014).
- [25] David W. Johnson and Roger T. Johnson, An educational psychology success story: Social interdependence theory and cooperative learning, *Educ. Res.* **38**, 365 (2009).
- [26] Bethany R. Wilcox and H. J. Lewandowski, Developing skills versus reinforcing concepts in physics labs: Insight from a survey of students' beliefs about experimental physics, *Phys. Rev. Phys. Educ. Res.* **13**, 010108 (2017).
- [27] Maura Borrego, Jennifer Karlin, Lisa D. McNair, and Kacey Beddoes, Team effectiveness theory from industrial and organizational psychology applied to engineering student project teams: A research review, *J. Engin. Educ.* **102**, 472 (2013).
- [28] Sonia M. Goltz, Amy B. Hietapelto, Roger W. Reinsch, and Sharon K. Tyrell, Teaching teamwork and problem solving concurrently, *J. Management Educ.* **32**, 541 (2008).
- [29] Simone Volet, Mark Summers, and Joanne Thurman, High-level co-regulation in collaborative learning: How does it emerge and how is it sustained?, *Learn. Instr.* **19**, 128 (2009).
- [30] Hanna Järvenoja and Sanna Järvelä, How students describe the sources of their emotional and motivational experiences during the learning process: A qualitative approach, *Learn. Instr.* **15**, 465 (2005).
- [31] Myung Hwa Koh and Janette R. Hill, Student perceptions of group work in an online course: Benefits and challenges,

- J. Distance Educ. Rev. Educ. Distance **23**, 69 (2009), <https://www.ijede.ca/index.php/jde/article/view/477>.
- [32] Alexandra Werth, Jessica R. Hoehn, Kristin Oliver, Michael F.J. Fox, and H.J. Lewandowski, Instructor perspectives on the emergency transition to remote instruction of physics labs [Phys. Rev. Phys. Educ. (to be published)].
- [33] Hanna Järvenoja, Simone Volet, and Sanna Järvelä, Regulation of emotions in socially challenging learning situations: An instrument to measure the adaptive and social nature of the regulation process, *Educ. Psychol.* **33**, 31 (2013).
- [34] Jessica L. Wildman, Daniel M. Nguyen, Ngoc S. Duong, and Catherine Warren, Student teamwork during COVID-19: Challenges, changes, and consequences, *Small Group Res.* **52**, 119 (2021).
- [35] Erin L. Dolan, Course-based undergraduate research experiences: Current knowledge and future directions, Natl. Res. Counc. Comm. Pap. **1**, 1 (2016), <https://advance.louisiana.edu/sites/advance/files/Course-based%20Undergraduate%20Research%20Experiences%20Current%20knowledge%20and%20future%20directions.pdf>.
- [36] Lisa A. Corwin, Mark J. Graham, and Erin L. Dolan, Modeling course-based undergraduate research experiences: An agenda for future research and evaluation, *CBE Life Sci. Educ.* **14**, es1 (2015).
- [37] Arundhati Bakshi, Lorelei E. Patrick, and E. William Wischusen, A framework for implementing course-based undergraduate research experiences (cures) in freshman biology labs, *Am. Biol. Teach.* **78**, 448 (2016).
- [38] Josh T. Beckham, S. Simmons, Gwendolyn M. Stovall, and James Farre, The freshman research initiative as a model for addressing shortages and disparities in STEM engagement, *Directions for Mathematics Research Experience for Undergraduates* (World Scientific, Singapore, 2015), pp. 181–221.
- [39] Rachel Kemp Alexander Chippendale, Monica Harrelson, Jennifer Shumway, Amanda Tan, Sarah Zuraw, and Jennifer L. Ross, Microscope project for undergraduate laboratories, [arXiv:1606.03052](https://arxiv.org/abs/1606.03052).
- [40] Rebecca L. Walcott, Phaedra S. Corso, Stacia E. Rodenbusch, and Erin L. Dolan, Benefit-cost analysis of undergraduate education programs: An example analysis of the freshman research initiative, *CBE Life Sci. Educ.* **17**, rml (2018).
- [41] Rory Waterman and Jen Heemstra, *Expanding the CURE Model: Course-based Undergraduate Research Experience*, Research Corporation for Science Advancement (2018), <https://rescorp.org/gdresources/publications/Expanding-the-CURE-Model.pdf>.
- [42] James Paul Masona, Courtney Peck, Alexandra Werth, Colin G. West, Donald Woodraska, H.J. Lewandowski *et al.*, Solar soft x-ray flare frequency distribution (to be published).
- [43] Leonard Springer, Mary Elizabeth Stanne, and Samuel S. Donovan, Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis, *Rev. Educ. Res.* **69**, 21 (1999).
- [44] Phyllis C. Blumenfeld, Ronald W. Marx, Elliot Soloway, and Joseph Krajcik, Learning with peers: From small group cooperation to collaborative communities, *Educ. Res.* **25**, 37 (1996).
- [45] Piet Van den Bossche, Wim H. Gijssels, Mien Segers, and Paul A. Kirschner, Social and cognitive factors driving teamwork in collaborative learning environments: Team learning beliefs and behaviors, *Small Group Res.* **37**, 490 (2006).
- [46] Simone E. Volet and Grace Ang, Culturally mixed groups on international campuses: An opportunity for intercultural learning, *Higher Educ. Res. Dev.* **17**, 5 (1998).
- [47] Lazarus Ndiku Makewa, Dorcas Gitonga, Baraka Ngussa, Samwel Njoroge, and Joshua Kuboja, Frustration factor in group collaborative learning experiences, *Am. J. Educ. Res.* **2**, 16 (2014).
- [48] Gavriel Salomon and Tamar Globerson, When teams do not function the way they ought to, *Int. J. Educ. Res.* **13**, 89 (1989).
- [49] Deborah L. Butler and Philip H. Winne, Feedback and self-regulated learning: A theoretical synthesis, *Rev. Educ. Res.* **65**, 245 (1995).
- [50] Ernesto Panadero and Sanna Järvelä, Socially shared regulation of learning: A review, *Eur. Psychologist* **20**, 190 (2015).
- [51] David Garratt-Reed, Lynne D. Roberts, and Brody Heritage, Grades, student satisfaction and retention in online and face-to-face introductory psychology units: A test of equivalency theory, *Front. Psychol.* **7**, 673 (2016).
- [52] Guan-Yu Lin, Scripts and mastery goal orientation in face-to-face versus computer-mediated collaborative learning: Influence on performance, affective and motivational outcomes, and social ability, *Comput. Educ.* **143**, 103691 (2020).
- [53] H. S. Hudson, Solar flares, microflares, nanoflares, and coronal heating, *Sol. Phys.* **133**, 357 (1991).
- [54] Norma B. Crosby, Markus J. Aschwanden, and Brian R. Dennis, Frequency distributions and correlations of solar x-ray flare parameters, *Sol. Phys.* **143**, 275 (1993).
- [55] Astrid Veronig, M. Temmer, A. Hanslmeier, W. Otruba, and M. Messerotti, Temporal aspects and frequency distributions of solar soft x-ray flares, *Astron. Astrophys.* **382**, 1070 (2002).
- [56] James Paul Mason, Raphael Attie, Charles N. Arge, Barbara Thompson, and Thomas N. Woods, The sdo/eve solar irradiance coronal dimming index catalog. i. methods and algorithms, *Astrophys. J. Suppl. Ser.* **244**, 13 (2019).
- [57] S. S. Grover, T. A. Ito, and B. Park, The effects of gender composition on women's experience in math work groups, *J. Pers. Soc. Psychol.* **112**, 877 (2017).
- [58] Katherine N. Quinn, Michelle M. Kelley, Kathryn L. McGill, Emily M. Smith, Zachary Whipps, and N. G. Holmes, Group roles in unstructured labs show inequitable gender divide, *Phys. Rev. Phys. Educ. Res.* **16**, 010129 (2020).
- [59] The teamwork training and scenarios were loosely based off the The Center for the Improvement of Mentored Experiences in Research (CIMER) training started at University of Wisconsin—Madison, <https://cimerproject.org/>.

- [60] Sanna Järvelä, Simone Volet, and Hanna Järvenoja, Research on motivation in collaborative learning: Moving beyond the cognitive-situative divide and combining individual and social processes, *Educ. Psychol.* **45**, 15 (2010).
- [61] Hanna Järvenoja and Sanna Järvelä, How students describe the sources of their emotional and motivational experiences during the learning process: A qualitative approach, *Learn. Instr.* **15**, 465 (2005).
- [62] Simone Volet and Caroline Mansfield, Group work at university: Significance of personal goals in the regulation strategies of students with positive and negative appraisals, *Higher Educ. Res. Dev.* **25**, 341 (2006).
- [63] This Likert scale was modified from the ranking scale on the original scale AIRE survey by Järvenoja *et al.*, which asked students to pick a number 1–4, where 1 was labeled not very important.
- [64] This was changed from the original AIRE survey, in which students were asked to rank their first most important goal, second most important goal, and least important goal, to cut back on the length of the survey due to time constraints of the student respondents.
- [65] Again, this scale has been modified from the original AIRE survey which asks students if “it did not happen at all.”
- [66] Stephen Kanim and Ximena C. Cid, Demographics of physics education research, *Phys. Rev. Phys. Educ. Res.* **16**, 020106 (2020).
- [67] Kilem Gwet, Kappa statistic is not satisfactory for assessing the extent of agreement between raters. Series: Statistical Methods For Inter-Rater Reliability Assessment, (1), 2002, <https://www.semanticscholar.org/paper/Kappa-Statistic-is-not-Satisfactory-for-Assessing-Gwet/7786ba48592a8b6ae773a8385a156154e02f4534>.
- [68] Henry I. Braun, Douglas N. Jackson, and David E. Wiley, Socially desirable responding: The evolution of a construct, *The Role of Constructs in Psychological and Educational Measurement* (Routledge, London, 2001), pp. 61–84.
- [69] F. Bu, A. Steptoe, and D. Fancourt, Who is lonely in lockdown? Cross-cohort analyses of predictors of loneliness before and during the COVID-19 pandemic, *Public Health* **186**, 31 (2020).
- [70] Rachel C. Conrad, Hyeouk “Chris” Hahm, Amanda Koire, Stephanie Pinder-Amaker, and Cindy H. Liu, College student mental health risks during the COVID-19 pandemic: Implications of campus relocation, *Journal Psychiatric Res.* **136**, 117 (2021).
- [71] B. Latané, K. Williams, and S. Harkins, Many hands make light the work: The causes and consequences of social loafing, *J. Personality Social Psychol.* **37**, 822 (1979).
- [72] Charles Crook, Motivation and the ecology of collaborative learning, *Rethinking Collaborative Learning* (Free Association Press, 2000), pp. 161–178.
- [73] Google LLC, Google Colaboratory faq page (2020).
- [74] A. Kelly Lane, Clara L. Meaders, J. Kenny Shuman, MacKenzie R. Stetzer, Erin L. Vinson, Brian A. Couch, Michelle K. Smith, and Marilyne Stains, Making a first impression: Exploring what instructors do and say on the first day of introductory STEM courses, *CBE Life Sci. Educ.* **20**, ar7 (2021).
- [75] Shannon B. Seidel and Kimberly D. Tanner, What if students revolt?—Considering student resistance: Origins, options, and opportunities for investigation, *CBE Life Sci. Educ.* **12**, 586 (2013).
- [76] Clara L. Meaders, Emma S. Toth, A. Kelly Lane, J. Kenny Shuman, Brian A. Couch, Marilyne Stains, MacKenzie R. Stetzer, Erin Vinson, and Michelle K. Smith, “What will i experience in my college stem courses?” An investigation of student predictions about instructional practices in introductory courses, *CBE Life Sci. Educ.* **18**, ar60 (2019).
- [77] Denise M. Anderson, Francis A. McGuire, and Lynne Cory, The first day: It happens only once, *Teach. Higher Educ.* **16**, 293 (2011).
- [78] Anthony D. Hermann, David A. Foster, and Erin E. Hardin, Does the first week of class matter? A quasi-experimental investigation of student satisfaction, *Teach. Psychol.* **37**, 79 (2010).
- [79] Jared J. McGinley and Brett D. Jones, A brief instructional intervention to increase students’ motivation on the first day of class, *Teach. Psychol.* **41**, 158 (2014).
- [80] Jonathan Osborne, Sue Collins, Mary Ratcliffe, Robin Millar, and Rick Duschl, What “ideas-about-science” should be taught in school science? A delphi study of the expert community, *J. Res. Sci. Teach.* **40**, 692 (2003).
- [81] Alexandra Werth, Kristin Oliver, Colin G. West, and H. J. Lewandowski, Engagement in collaboration and teamwork using google colaboratory, in *Proceedings of the PER Conference* (2022), <https://www.compadre.org/Repository/document/ServeFile.cfm?ID=16280&DocID=5649>.