Evaluating the role of student preference in physics lab group equity

N. G. Holmes [1,*] Grace Heath [2] Katelynn Hubenig,[1,3] Sophia Jeon,[1] Z. Yasemin Kalender,[1]
Emily Stump [1] and Eleanor C. Sayre [3]

[1]*Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14853, USA*
[2]*Loyola University New Orleans, New Orleans, Louisiana 70118, USA*
[3]*Department of Physics, Kansas State University, Manhattan, Kansas 66506, USA*

Physics education research is replete with observations and proposed explanations for gender disparities in physics. In this work, we operationalize a definition for equity as everyone has access to the learning environment and everyone's voice is heard (adapted from previous definitions). We review prior research that observed inequities in physics lab group work and evaluate the degree to which these inequities may arise from student preferences. Regarding access to the learning environment, we find that men and women have similar preferences for experimentation roles, though women generally preferred sharing, rather than dividing or rotating, roles. Regarding everyone's voice being heard, we find that students prefer when no single student takes charge, though many preferred that students take turns leading the group, with no significant differences between men's and women's preferences. We also find no significant differences in men's and women's preferences for the gender composition of their group. We conclude that the observed inequities are not explained by student preference and explore implications for instruction and future research.

I. INTRODUCTION

This paper aims to take a multipronged approach to understanding equity in undergraduate physics lab instruction. Physics remains one of the least diverse disciplines in science, technology, engineering, and math (STEM) with little evidence of improvements over the last decade [1]. This stagnation in representation persists despite significant improvements in the diversity of students attending higher education [2] and majoring in other STEM disciplines, such as biology and engineering [1]. Researchers have demonstrated that historically excluded and minoritized (HEM) students and women often receive similar course grades as majority students [3–6]. Students, however, choose their major based on many variables beyond their course performance [7]. Students' perceptions of the disciplines and their perceptions of succeeding in the discipline seem to be a critical factor in their choice. Indeed, HEM students and women in physics report self-efficacy and physics identity levels that do not match their course performance [3,8–12].

One variable that connects to students' perceptions of a discipline is the students' sense of community and belonging [13,14]. The interactions with instructors and other students significantly influence those perceptions. For example, how much students interact with their peers in and out of class positively impacts their engagement, motivation, and science identity [15–18] and their likelihood to persist in the discipline [17,19,20]. While a few studies in the life sciences have found that increased interactivity during class particularly improves HEM students' and women's performance [21–26], many studies in physics have failed to see the same effects [4,27,28]. One possible explanation for the inconsistency may relate to the details of implementation. Each student must be actively engaged for interactive instruction to be effective for all students. If the interactions between peers are inequitable, then one student may be doing all the active thinking, while the others participate only passively. For example, in lecture courses, researchers have found that, without deliberate intervention, women participate considerably little in discussions, asking and answering questions, or presenting material [29–33]. Instead, students with high confidence or self-efficacy may dominate the group work and thus may be the ones benefiting from the interactivity. Unfortunately, there has been limited, and often conflicting, work understanding group dynamics in physics education specifically [34].

In this work, we examine students' experiences in physics lab group work with a focus on gender. Instructional labs

*ngholmes@cornell.edu

offer an ideal context to study group work equity because students generally work in small groups for extended periods of time on a shared task (in contrast to group work in lectures or tutorials, for example, where students submit individual work in consultation with their group mates). Furthermore, many studies in physics education (described in more detail below) have found inequities in how students engage in lab group work. The goal of the paper is to answer the question: How do gender inequities observed in lab work relate to students' group work preferences? By understanding the role of student preferences, we can then inform instructional strategies that accommodate preference, while attempting to mitigate inequities.

II. OPERATIONALIZING GENDER EQUITY IN LAB WORK

We focus here on interactions between lab group members and the learning environment, working from a definition that equity means *everyone has access to all parts of the learning environment and every voice is heard.* This definition builds from previous work that has defined equity as "the fair distribution of opportunities to learn" [35], "all students have equitable opportunities to learn" [36], "each group member has affordances to participate" [37], and "each group member feels comfortable sharing with their group, and their voice is heard by all group members" [38] (in this last formulation and our definition, "voice" and "heard" are taken metaphorically). We specifically focus our definition on versions of equity related to the activities associated with group work in labs, explicitly avoiding discussion about equity of learning [or other, quantitative models of equity related to student outcomes; see, e.g., Refs. [39–41]]. With this definition, we follow guidance from Carlone *et al.* [42] to consider equity associated with the culture of the classroom rather than with measures of student outcomes.

A. Equal access to the learning environment

In a lab class, the first component of our definition (that *everyone has access to the learning environment*) relates to students' access to handle the equipment, analyze the data, and document and communicate the results. Many labs have explicit aims for students to develop these skills technically [43]. From a participationist perspective of learning, we would assume that actively participating in the learning activity is important for student learning. In labs, we assume that students will, for example, learn more about data analysis by analyzing the data or more about experimental design by setting up the equipment and collecting data.

Through this participation, lab courses are opportune for developing a community of practice [44] and students' attitudes towards the discipline [45,46]. However, many labs instead negatively impact students' attitudes towards

physics [45,46] and disciplinary identities [47] for a variety of reasons. For women in particular, participation in lab may be problematic due to the ubiquitous group work [48].

Observations of students in labs have consistently found inequitable divisions of labor when students conduct experiments [49–53]. In these studies, men and women systematically take on different roles. Often, men handle the equipment [49,51–53] or analyze data [50] more often than women, while women may take on managerial or note-taking roles [49–51]. Danielsson described these role negotiations as students "doing gender" while they navigate "doing physics" [54–56]. That is, students are simultaneously deciding how to proceed with the investigation while also deciding who should complete each task. In mixed-gender groups, gender-based stereotypes or self-efficacy gaps favoring men can (implicitly or explicitly) impact students' negotiations around role divisions, leading to inequitable task divisions [57,58]. Additional evidence towards this negotiation of gender in student lab roles comes from comparing the participation of students in mixed-gender versus single-gender groups. In our previous work [51,52], we found that men in single-gender groups were disproportionately high equipment users and women in mixed-gender groups were disproportionately high laptop users. The small number of all-women groups, however, left conclusions tentative, though it appeared their behavior did not differ from that of men in mixed-gender groups.

The negotiation of doing gender and doing physics suggests that there are manifestations of students' gender that result in students taking on one role over another. One possible explanation for this manifestation that we explore here is that men and women have different preferences for the lab roles. That is, could this negotiation just be the result of women preferring some roles more than men (and vice versa)? The observed participation of women in single-gender groups suggests not, but the data are too small to be conclusive.

B. Everyone's voice is heard

The second component of our definition (that *everyone's voice is heard*) relates to a constructivist perspective that externalizing thought processes to others is critical for everyone's learning [59,60]. Research has shown that interacting with others improves learning beyond individual active learning [60–64] and students who talk more in group work learn more [65]. Having everyone's ideas heard also helps the rest of the group learn [60].

However, research has found multiple examples of inequities in women's voices being heard. For example, men have been found to voluntarily respond to disproportionately more instructor questions in lecture [29,33] or to speak longer during group presentations [31] than women. One study found that women were overall more

uncomfortable participating in whole-class discussions than men [66]. Fortunately, strategies that remove barriers to speaking up in class (such as employing random call or modeling counterstereotypic group behavior) can remove or reduce the biased participation rates [31,33] and biased perceptions of women's abilities [67].

In labs, particularly ones where students engage in experimental design and critical thinking, being heard relates to guiding the process and outcome of the investigation. The students must make various decisions about how to conduct the investigation, which relates to their engagement in the cognitive tasks of experimentation [68–70]. In lab group work, therefore, students can contribute to the discussion through *directive bids*: statements that direct the group progress [71]. Students' voices are heard when they make and acknowledge each other's bids.

However, students can interpret directive bids differently depending on the issuer's gender. In a case study of two students in high school algebra taking on leadership positions in a group [72], the female student's directive bids were significantly more often ignored by her male peers and deemed inappropriate. In contrast, the male student's directive bids were successfully taken up, positioning him as central and helpful. Repeat experiences with unacknowledged bids seem to diminish the students' contributions overall, further marginalizing their voice being heard [38,72].

In our previous work [73,74], we evaluated the bid exchanges of eight groups of students during labs: four groups with equitable bid exchanges and four groups with inequitable bid exchanges. We found that most groups (whether equitable or inequitable) had a single student most in charge, from which we inferred that students need not be equally in charge (or have no one in charge) for the group to be equitable. This was in contrast to previous work that inferred imbalanced inchargeness led to inequitable group work [38,75]. In all four of our equitable groups, however, the group leaders actively used their inchargeness to involve the other group members and make sure their voices were heard. In contrast, leaders in our inequitable groups did not acknowledge others' bids or they dominated the bid exchanges. Importantly, the group's gender composition did not necessarily dictate the dynamics in these bid exchanges, contrary to existing literature [76,77]. That is, we found both equitable and inequitable groups made up of mixed-gender pairs, majority women trios, and majority men trios. In the trios, men and women were also equally likely to have equitable or inequitable bid exchanges with other men and women (that is, there was no pattern for higher equity between same-gender pairs). The results did, however, suggest a relationship between gender, inchargeness, and equity: three of the four equitable groups had a female student most in charge and three of the four inequitable groups had a male student most in charge.

This relationship to gender may be a result of men and women having different preferences for and experiences with small group work [66]. For example, women in single-gender groups were found to interact more and to provide more information than women in mixed-gender groups when solving cooperative problems in physics [78]. Their problem-solving behaviors also differed, with women in single-gender groups spending more time planning their solution strategies than women in mixed-gender groups. In all cases, women in mixed-gender groups similarly differed in their interactions and problem solving as their male partners. These differences in interactions have been partially associated with men's and women's preferences for interaction roles in group work [66]. In that study, a higher proportion of women preferred a collaborative role (as opposed to a leader, listener, or recorder), while more men preferred a leadership role.

In the current work, therefore, we evaluate whether men and women differ in their preferences for their group's leadership structures (separate from the particular role they prefer to take on themselves). These different preferences may lead men and women to perceive their own and others' inchargeness in different ways, leading to the observed inequities in students' voices being heard.

C. Limitations

In this work, we explore issues of equity related specifically to gender. All students self-identified their gender on a research survey, with approximately 20%–30% of students identifying as women across the cohorts. The number of students who identified as nonbinary were too small to include in the study while protecting their anonymity. We acknowledge the significant limitation of the binary treatment of gender in our data [34].

The data we reference are also limited in that they come from a population of students unfortunately too homogeneous in terms of race, ethnicity, and other demographic or identity characteristics for us to explore other forms of (in)equity. The students in the sample were primarily Caucasian or white or Asian or Asian-American, with proportions not representative of the institution's student population [79] or population of physics students [80,81]. We acknowledge this limitation particularly in light of the much larger conversation in physics education research about equity in a variety of forms.

Important research has explored intersections of, for example, race, culture, gender, and physics identity [82–88]. Critical theory (including critical gender theory and critical race theory) acknowledges the complexity of equity along axes that are not simply binary gender. Furthermore, they challenge traditional structures (such as of power and privilege) that reproduce inequities in education [24,84,89]. While our data contribute to this conversation, the focus on gender alone ignores important dimensions of students' identities that impact their

experiences in lab group work [34]. In the course of the analysis, we find that the associated dynamics in group work were too complicated to concern gender alone. Future work should evaluate the ways in which the inferences made here may translate to group dynamics based on other dimensions, including race, ethnicity, and other identity characteristics.

We also acknowledge that the population from which our data are drawn is not representative of physics majors or students taking physics courses nationally or internationally [90]. The students are highly motivated to engage in physics and so their behaviors may not generalize to other samples (such as nonphysics majors). The proportion of women in the class, however, was representative of national surveys of physics students [1,81]. Thus, any interpretations based on proportions of men and women in the classroom may be appropriate for similarly represented classrooms. While many of our results reflect those of international studies [50,52,54–56,91], the culture of the physics lab and physics community in this sample may also not be representative of students in other countries.

Survey responses analyzed in the study were also collected at the start of the mechanics course during the shift to remote instruction as a response to the COVID-19 pandemic. The students' perspectives reflect their experience *before* taking college-level physics, which would primarily include in-person experiences during high school. Thus, we do not expect the shift to remote instruction to have significantly impacted their perspectives, though this should not be ignored when interpreting the results.

Finally, the interviews included only a subset of the students and so may not be representative of the class as a whole. The demographics of the interview sample were representative of the class as a whole, but this meant that the voices of women were less represented in the data (based on the representation of women in the class).

III. METHODS

A. Research context

Data were collected from undergraduate students at Cornell University enrolled in the honors-level mechanics course of a calculus-based physics sequence between 2019 and 2020. Approximately 70% of the students were physics majors, with the remainder either undeclared or majoring in another science or engineering field. Approximately 30% of the students self-identified as women and 70% as men. Our data come from two sources: interviews and surveys. The interviews were conducted midway through the Fall 2019 semester (typical in-person instruction). The surveys were distributed to students at the start of the course in Fall 2020 (during emergency remote teaching due to the COVID-19 pandemic).

The labs for this course were designed to emphasize the process of experimentation in physics (see, for example,

Refs. [46,70,92,93]). Students were expected to attend labs approximately weekly (with some weeks off due to holidays or exams) and each lab session was two hours long. Each experiment spanned two or three lab sessions. As discussed elsewhere, students were increasingly responsible for designing their investigations, with a fully open-ended project lab at the end of the semester [94]. Students worked collaboratively in groups of two or three students to design and implement their experiments and submitted one electronic notebook as a group to be graded.

Another aspect of the research context is our positionality as authors. Authors on this paper identify as cis-women and nonbinary, identify as physicists, and our racial identities include white and Asian. At the time of performing this research, our professional appointments included undergraduate students, a graduate student, a postdoc, a pre-tenure faculty member, and a full professor. While the data represented here are from a U.S. institution, our citizenship is more varied, including the U.S., Canada, and Turkey.

B. Interviews

We conducted individual, semistructured interviews about students' experiences in the lab course. Students were recruited through an in-class announcement and submitted their availability for an interview. A subset of the interested participants were contacted for interviews based on their availability and to obtain a diverse group of students with diverse perspectives. We considered students' gender, major, and lab section in selecting a diverse group of participants. Of 89 students in the class, 24 students volunteered for interviews, of which we scheduled 14 (six women and eight men, self-identified). Three of the scheduled women did not show up to the interview, so our final sample consists of 11 students (three women and eight men). The gender composition of the sample is representative of the class as a whole (30% women).

We developed the semistructured interview protocol focusing on three broad themes: students' perception of labs as an authentic research experience, understanding students' experiences working in groups, and probing the lab's influence on students' perceptions of physics. We conducted pilot interviews with six students (five upper-division undergraduate physics majors and one graduate physics student) and made only minor modifications to the interview protocol following the pilots. Two researchers conducted all interviews separately. The interviews were audio recorded and the researcher took notes during the interview to supplement the audio. Each interview took approximately one hour. At the end of the session, students were compensated for their time with a $15 Amazon gift card.

Three researchers generated transcripts of the interviews and coded them for several overarching themes. Here we focus only on the statements related to group work.

Two different researchers then evaluated the statements identified as group work and further coded them according to whether they related to statements about leadership and about roles in the lab.

Statements about roles were divided into two categories: Experience with (or preference for) consistent roles and experience with (or preference for) rotating roles. Statements categorized for consistent roles indicated the student experienced or preferred each group member take on a single role at a time. For example, "I've done the experimentation at times if someone wants to do lab notes." In some cases, these statements included concessions to take on undesirable roles out of compromise. Statements categorized for rotating roles indicated that the student had no preference for or experience with a particular role and saw the benefit of having a well-rounded experience, such as by rotating or switching roles. For example, "We've been dividing the work so that in some sessions, I've been actually doing the experiment. I've been collecting and analyzing data. I've been explaining the lab notes. I got a taste of all three roles, without the pressure of having to do all of them in a single period."

Researchers divided statements coded for leadership into two categories as well: experience with (or preference for) hierarchy and experience with (or preference for) no hierarchy. Statements categorized for hierarchy indicated the student saw value in having a leader in the group, preferred that someone take on a leadership position, or indicated a norm for having a leader in the group. For example, "I think it could be helpful if it wasn't the same person in charge every time, like if everyone got the experience what it is like to be in charge that would be helpful." In these statements, students typically acknowledged the benefit of having a leader and the positive impacts of a leader on the group dynamic. We did not include statements regarding hierarchies or power dynamics between students and the teaching assistant because they did not relate to the leadership structure within the group. Statements categorized for no hierarchy referred to an explicit mention of a preference for or experience with no leadership in the group or everyone being equal. These statements typically indicated that the student valued equal contributions from all group members and open collaboration. For example, "What I like about [my lab partners]—everybody has equal work, everyone is willing to try each other's ideas, nobody's taking charge."

C. Surveys

Inspired by the interviews, we developed a survey to broadly probe students' perceptions of group work at the start of the semester. Students' responses were used to intentionally form lab groups, so all students were required to take the survey as part of their course, though students were given the option to exclude their responses from the research. In total, 89 students consented to complete the survey (100% of students enrolled that semester), 26 who identified as women, 60 who identified as men, and three who identified as nonbinary or did not identify their gender.

We analyzed two survey items about students' preferences for lab roles. The first question asked "Which of the following experiment tasks do you prefer taking on?"
 (a) Setting up the apparatus and collecting data.
 (b) Writing up the lab procedures and conclusions.
 (c) Analyzing data and making graphs.
 (d) Managing the group progress.
 (e) No preference or none of the above.
The answer choices were chosen to align with lab roles identified in previous work [50,51] and respondents could select multiple options. The second question asked "Which of the following approaches to group tasks do you prefer?"
 (a) One where each person has a different task.
 (b) One where everyone works on each task together.
 (c) One where everyone takes turns with each task.
 (d) No preference.
 (e) Something else.
The answer choices were based on the student interviews. For this question, respondents could only select one answer, with the option to write in their own answer for (e) if they did not agree with any of the choices provided.

We also analyzed two survey items related to everyone's voices being heard. The first item, about students' leadership preferences within a group, asked "Which of the following approaches to leadership do you prefer?"
 (a) One where one student regularly takes on the leadership role.
 (b) One where no one takes on the leadership role.
 (c) One where the leadership role rotates between students.
 (d) No preference.
 (e) Something else.
Respondents could only select one response, with the option to write in their own answer to (e) if they did not agree with any of the choices provided. The second item asked "What gender breakdown of group do you prefer to work in?"
 (a) Majority female identifying.
 (b) Majority male identifying.
 (c) No preference.
 (d) Other.
Respondents could only select one response, with the option to write in their own answer to (d). Four students wrote in an answer that indicated preference for an equal or diverse balance of female- and male-identifying students.

For each survey question, we calculated the proportion of men and women who selected each option.

IV. RESULTS

We present the results of the interviews and surveys split by the two forms of equity.
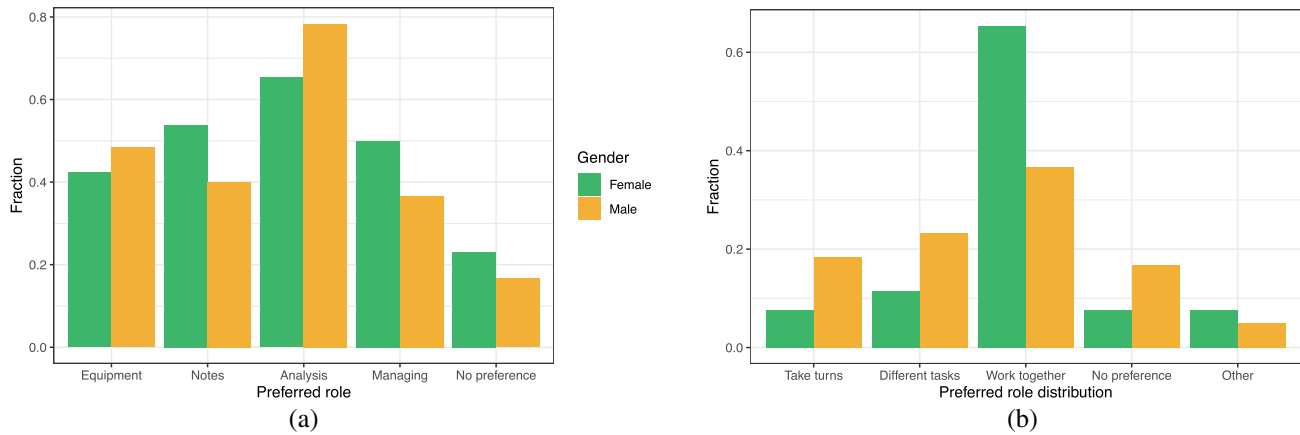
FIG. 1. Student responses to survey items about (a) preferred lab roles and (b) preferred group distribution of roles. Similar proportions of men and women preferred handling equipment, while 10% more women than men preferred taking lab notes or managing the group, while 15% more men than women preferred analyzing data. Most students (though a much larger fraction of women) preferred working together in each task.

A. Equal access to the learning environment

1. Interview data

In the interviews, all students described that students typically take on individual, consistent roles in the lab (Table I), either taking notes, collecting data and handling equipment, or analyzing data. Eight of the 11 students also identified benefits of rotating these roles either within or between lab sessions. For example,

"The issue with the idea of assigning specific roles…[is that it] may be less beneficial for the whole group. Because if one person is doing lab notes for the whole time then they are not engaged with the lab… A lot of people are fine with switching up to a tedious role from an active role if they [had done] the active role the previous week."

In contrast, some students identified issues with switching roles:

TABLE I. Number of statements coded for the two categories related to group work roles. All genders were self-reported by the students.

		Number of statements	
Student	Gender	Consistent roles	Rotating roles
S1	M	3	3
S2	M	1	4
S3	M	6	1
S4	M	4	1
S5	M	1	1
S6	M	2	1
S7	M	7	0
S8	M	2	0
S9	F	4	0
S10	F	13	3
S11	F	3	3

"For example if we keep switching the role there will be more mistakes discovered in the lab report and—well—different people will not develop in the way that will include different people's ideas."

Other students indicated that some roles (such as writing notes) should be divided while others (such as collecting data) should be shared.

Students were quite vague, however, in describing how roles were assigned. Most comments indicated informal assignments, consistent with previous observations [51]:

"I don't feel like we explicitly talk about it—it just comes about. Someone just starts building the experiment, someone just happens to have Excel open. Sometimes people volunteer so you let them do it. Otherwise it arises naturally."

A couple of interviewees indicated students stepping in to fill empty roles, two students mentioned roles being assigned based on student ability, and one described play "nose goes" for who had to take on lab notes (an undesirable role in their eyes).

From the interviews, we get a sense of collegiality among the students and a fairness in access to the learning environment. There were no clear patterns in preferences based on gender, though only three of the interviewees were women. We turn to the role preferences survey to explore these preferences with a larger sample size.

2. Survey data

Figure 1(a) shows the proportion of men and women who indicated a preference for each lab role. The proportion of men and women preferring to set up the apparatus and collect data are similar. A slightly larger proportion of woman than men indicated a preference for writing the lab notes and managing the group, while a slightly larger

proportion of men than women indicated a preference for analyzing data.

This observation is an interesting contrast to previous work observing students in labs, described above [51]. In that work, we saw that women in mixed-gender groups spent less time on the apparatus than women in single-gender groups or men. Women spent as much (or more) time on the equipment than men (in any type of group) and similar amounts of time with the other activities when working only with other women. At face-value, these observations are unsurprising. In any group, students must negotiate roles and divide tasks between themselves and over time. In any lab activity, *someone* must handle the equipment, analyze the data, and so on. In all-women groups, that *someone* must be a woman. However, the mixed-gender groups indicate that some roles are negotiated with gender in mind [54,55]. Our survey data, however, add nuance to this interpretation. Women have similar preference for interacting with the equipment as men. More women, however, preferred taking notes, and so may have been more willing to step in to fill this role.

The implications of these inequities can be further inferred by students' preferences for dividing or sharing roles. Figure 1(b) shows the proportion of men and women who indicated a preference for each role sharing option. Here we see larger differences between the two genders. Women were much more likely to express a preference for working together with their group members on all tasks, while men were more evenly split among all of the options available. Of the students who chose to write in their own answers (Other), all suggested some combination of rotating tasks, working together on each task, and taking turns with tasks. This result is in stark contrast to the interview data, where the students generally indicated that roles would be divided or rotated—no interviewees indicated preference for (or experience with) working together on a role. The survey data suggest women may be particularly negatively impacted in mixed-gender groups based on their preference to work together on tasks.

B. Every voice is heard

1. Interview data

The interviewees generally preferred no hierarchy among the group members (i.e., no designated leader) with no patterns between gender (Table II). The only students who had zero statements related to no hierarchy had zero statements coded as leadership overall. Overwhelmingly, the statements related to no hierarchy referred to students not wanting anyone to be taking over or "bossing people around." For example, one student explicitly mentioned that he disagreed with the lab manual's instructions to assign a leader to direct the group:

"I don't know. I'm pretty happy with my lab partners so far. What I like about them? Everybody has equal work,

TABLE II. Number of statements by each interviewee coded for the two categories related to leadership in group work. All genders were self-reported by the students.

Student	Gender	Number of statements	
		Hierarchy	No hierarchy
S1	M	0	2
S2	M	1	8
S3	M	5	3
S4	M	0	2
S5	M	1	3
S6	M	0	0
S7	M	0	1
S8	M	0	0
S9	F	2	1
S10	F	0	0
S11	F	1	5

everyone is willing to try each other's ideas. Nobody's taking charge. Although in the lab book, it says that one person should take charge but I don't really like that."

One student indicated that leaders were not necessary if everyone got along. Several students explicitly alluded to the notion of communication and how equality helps having every voice heard. For example,

"I think working in a group is very helpful, you can bounce your ideas off of each other. If one person doesn't know a thing then another person can contribute. In my lab experience, everybody has been participating equally. There hasn't been an imbalance."

Some students, however, also acknowledged the possible value of having group leaders. For example, one student described that the need for leaders depended on whether everyone was contributing ideas:

"I guess in this lab, there is not a hierarchy I would say. Maybe the person who comes up with the idea might be leading the group but if everybody agrees on it and everyone has an idea of what to do, then there is no leader. But if one person comes up with an idea and the other people don't have an idea, then yeah there is a hierarchy there."

All of the students that mentioned a preference for hierarchy also indicated preference for no hierarchy, reinforcing the notion that there may be a time and place for one or the other. Once again, there was no clear pattern in preferences based on gender.

2. Survey data

Figure 2 shows the proportion of men and women who indicated a preference for each group leadership option.
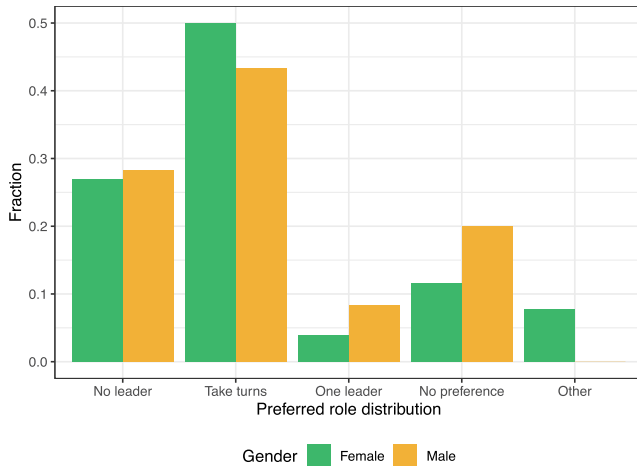
FIG. 2. The fraction of men and women with each leadership preference. Most students preferred taking turns as leader with only small differences between men and women's preferences.

Consistent with the interviews, very few students indicated a preference for a single leader in the group. Inconsistent with the interviews, most students preferred students take turns as the leader. A slightly higher proportion of women than men expressed a preference that group members take turns acting as group leader, while a slightly higher proportion of men than women indicated no preference. The latter may reflect that men are somewhat less concerned that a leader may negatively impact their lab experience. Only women chose the option to write in their own response. Two of them expressed the opinion with qualification that they prefer sharing the leadership positions, while the other favored a single leader so long as that person listened well to others.

We also found that men and women overwhelmingly had no preference for their group's gender composition (78% and 73%, respectively). A slightly higher proportion of women than men indicated a preference for a majority female-identifying group (19% and 2%, respectively) and a slightly higher proportion of men than women indicated a preference for a majority male-identifying group (13% and 8%, respectively). The implications of these proportions are discussed below.

V. DISCUSSION

In this paper, we defined lab group equity as everyone having access to the learning environment and everyone's voice being heard. The overarching conclusion from this work is that gender (in)equities in physics lab work are not simply due to student preference.

A. Summary of equity as access to the learning environment

We examined access to the learning environment through the experimentation roles that students prefer, as well as how those roles are divided or shared. In previous work, we and other researchers found that access to lab roles differs for men and women [49,51,53–56] and for students in single- and mixed-gender groups [51,52]. We infer that students in mixed-gender groups must navigate gender dynamics while also navigating the experiment—so-called "doing gender" while "doing physics" as described by Danielsson and colleagues [54,55]. In single-gender groups, women need only navigate the experiment—they can freely "do physics."

The interview and survey data support this interpretation. Women were equally interested in working with the experimental apparatus as men, suggesting that the inequities previously observed are not due to student preference. Instead, the interviews illuminated that students implicitly negotiated or fell into roles—negotiations that could be prone to implicit biases, such as perceptions that women excel in managerial or passive roles [49,57,91,95], are less associated with science than men [96–98], or are less competent in scientific roles than men [99]. Thus, biases, stereotypes, or microaggressions may lead to the gender dynamics in task negotiation when there is no explicit discussion of students' roles. Rarely, interviewees indicated roles being assigned based on student ability. Whether or not men and women differ in their equipment-handling, note-taking, or data-analyzing abilities, research suggests students' perceptions of these abilities may be biased [6,67]. Women's slight preference for taking notes may also explain why, in mixed-gender groups, women were found to spend more time on the laptop than men (students had to submit electronic lab notes for the group to be graded). One interviewee described playing nose goes when deciding who would take the lab notes–an undesirable role in his view. We might expect women to be less concerned about stepping in to take on this task than men in such a negotiation because fewer women than men found this role undesirable.

Men and women, however, differed in their preferences for how roles are shared or divided. Both men and women had a general preference for working together and sharing roles rather than dividing tasks, despite clear signs that students divide tasks according to the interviews and previous work [51]. Women's preference for working together was overwhelming, however. Given the inequities observed in student roles in mixed-gender groups, this strong preference for sharing roles may be a result of women's prior experiences not having access to various roles.

Throughout our analysis and interpretations, we implicitly infer that imbalanced access to the learning environment (through experimentation roles) is undesirable. Through our participationist lens, and because hands-on experience can be important for physics students' future career aspirations [100], a systematic lack of experience with the various aspects of experimentation is likely to

disadvantage the student. This applies to both men and women. With men spending more time with the equipment or data analysis, they are missing out on opportunities to develop scientific communication skills through writing the lab notes. With women spending less time with the equipment or data analysis, they are missing out on opportunities to develop important hands-on and analytic skills. This study importantly demonstrates that men and women are missing out on these opportunities despite similar preferences for engaging in the tasks. Students may then be systematically dissuaded from participating in these activities in the future, further widening the divide. Future work should evaluate whether students' preferences for these roles change over time as a result of continued limited access.

B. Summary of everyone's voice being heard

We examined being heard as the second component of equity through the preferences for leadership structures in student groups. In previous work, we found evidence of gender dynamics in the relationship between inchargeness and this form of group equity [73,74]. Namely, mixed-gender groups were most often equitable when a woman was most in charge. This observation is in direct contrast to the interviewed students' overwhelming preference for no one to take charge and all students to be equal, which in turn conflicts with the surveyed students' overwhelming preference for a single but rotating leader.

While these observations seem to be in conflict, the different perspectives can help to paint a coherent story. For example, in the interviews, the students typically interpreted the notion of a person in charge as being a "boss" or "taking over." Thus, students may have been referring to the kind of inchargeness that would lead to inequitable group interactions, such as that seen in Refs. [38,73–75]. In addition, many of the interviewees also acknowledged the potential benefits of having a leader in the group, particularly depending on the overall group dynamics or composition. This vacillating may help explain the survey results, where the most students (over 40%) preferred group members taking turns as leader. This choice relates to the interviewees' desires for everyone to be treated equally, while also relating to our previous observations of the potential benefits of a group leader in the equitable groups [73]. The second most popular preference, however, was to have no leader, which is consistent with the interviews and prior work.

Gender, surprisingly, did not emerge as a clear factor in the data. Research on leadership structures suggests that other aspects of the group or individuals (such as group extraversion [58], women's responses to stereotype or identity threat [57], or students' overall social friend network [101]) may mediate any observed gender dynamics with inchargeness. Alternatively, additional identity characteristics outside of gender may be at play.

Gender was also not a significant factor in the survey item probing students' preferred group gender composition. Previous work suggests that women should be equally or overrepresented in a group for the group to be equitable in terms of having voices heard [76,77]. The survey item asking about students' preferred group gender composition, however, showed most women did not prefer to be in majority female-identifying groups and generally had no preference for the gender composition. Interventions to form groups based on gender identity, therefore, may be overly heavy-handed and not reflective of students' actual preferences. Once again, research should evaluate whether these preferences change as students build up experiences in groups with different gender compositions.

C. Implications for research and teaching

This work sought to evaluate the ways in which student preference may relate to the inequities observed in student lab work. In doing so, we identified several ways in which students' perspectives were in conflict with previous observations. For example, students were generally very positive about their group work experiences, despite common observations of group inequity during lab work. While students could identify the behaviors of problematic individuals (nonideal lab partners), they did not generalize that experience to larger trends about their peers. As another example, students indicated preferences for no leader (in the interviews) or rotating leaders (in the survey), despite observations that a single leader can support group equity. From these conflicts, we highlight that researchers must ask: whose experience is centered by your data source and what interpretation can we reasonably draw from the experience that is centered?

As researchers, we see many students' experiences at once and thus see patterns and trends. Students see only their own experiences, but they feel and experience them directly. One implication could be that researchers' observed trends are irrelevant unless students perceive the trends themselves. Alternatively, students may simply need to accumulate their own data through a larger set of experiences to observe the trend that researchers see (recall that the data were all collected from a first year, first semester physics course). As such, researchers may be predicting the future and, by acting on the observations of the researcher, we can intervene before the students have accumulated too many negative experiences. Additional observations and data on students' perspectives on group work over longer periods of time would help illuminate these possible hypotheses.

Alternatively, the disagreements between the survey data and other perspectives (interviews and previous work) may be due to when the data were collected. The survey data were collected at the start of a remote semester during the COVID-19 pandemic. We are skeptical that this context impacted the results significantly for two reasons. First, the

survey data were collected at the start of students' first college-level physics course, so their perspectives were built from their experiences in high school, which took place mostly before the pandemic. Second, aspects of the survey results agree with other forms of evidence collected in person, suggesting the survey data are not entirely systematically different. Future work should evaluate this explicitly, however.

Lastly, we acknowledge that conflicts between the interview data and other perspectives may be due to the small sample size, particularly the few women represented in those data. While the demographics of the interviewees were representative of the class as a whole, we should be hesitant in generalizing to the population from this limited sample. Instead, we can focus the interpretations of the interviews as examples of the range of student thinking, rather than on proportional representation (i.e., case-oriented rather than recurrence-oriented interpretations [102]).

With regards to instruction, we see overwhelmingly that inequities observed in student access to the learning environment and being heard are not systematically due to student preferences. Likely, the inequities are instead due to structural aspects of the learning environment. Thus, instructors must institute policies and practices to push against gender norms and stereotypes. Our work here, however, also indicates that attending to student perceptions may help students take up instructional strategies intended to support equity. For example, one of the interviewees explicitly called out a disagreement with a statement about group dynamics in the lab manual. The lab manual encouraged students to rotate between three roles (adapted from those in the cooperative learning literature [103,104]): principal investigator, communicator, and reviewer No. 2. The principal investigator was described as "Manages the group by ensuring everybody is on task, keeping the investigation on track to answer the research question, and creating room for everybody in the conversation." Despite the intentional alignment between our definition of equity through everyone being heard, students' perceptions of "one person managing the group" was a person who would "take charge" or "boss everyone around"—clearly an undesirable and inequitable hierarchy. Given the possible benefits of leaders in the group, we posit that alternative framings of the leadership role (such as a person who makes sure everyone's ideas are heard) may better align with students' preferences. In addition, instructors could explicitly address the issues with power

dynamics in group work (given that students were generally in consensus regarding the potential negatives of leadership) and emphasize that each group member (not just the leader) is responsible for ensuring that all students' voices are heard and everyone has access to the learning environment.

Finally, instruction should respond to the observation that students do not always perceive group work experiences as negatively as the research trends may suggest. As such, heavy-handed interventions (such as to assign roles or form groups based on identity characteristics) might be unnecessary or overly prescriptive. For example, a common recommendation is to not isolate minority students such that the groups are either gender balanced, majority women, or all men [76,77]. Our data indicate, however, that students do not have strong preferences for the gender breakdown of their group, suggesting that intentional group forming may not support their preferences. Instead, instructors could consider asking students in an anonymous presurvey what group gender composition they would prefer and setting up the groups accordingly.

Other recommendations in the literature are to have each student rotate through the various group roles [103]. Our data suggest students, particularly women, may prefer working on roles together and that some students may prefer taking on roles for which they are comfortable or confident. Based on the evidence, one-size-fits-all instructional approaches are unlikely to be better for all students.

Regardless, the work here and elsewhere indicates that structural change is necessary to promote equity in group work [25,77,105–109]. Future work should systematically test structural interventions, such as varied group structures and role assignments or forming groups based on student preference, to evaluate their impacts on both aspects of group equity, as well as students' perceptions of equity, to understand the possible tensions and solutions. Additionally, as we consider the expansion of this work to classes beyond ones populated primarily by physics majors and ones at primarily white institutions, we expect to see additional effects of race, major, and course culture.

ACKNOWLEDGMENTS

[1] Anne Marie Porter and Rachel Ivie, *Women in Physics and Astronomy, 2019*, Technical Report (American Institute of Physics, New York, 2019).

[2] Richard Fry, *U.S. Women near Milestone in the College-Educated Labor Force*, Technical Report (Pew Research Center, Washington, DC, 2019).

[3] Shima Salehi, Eric Burkholder, G. Peter Lepage, Steven Pollock, and Carl Wieman, Demographic gaps or preparation gaps?: The large impact of incoming preparation on performance of students in introductory physics, Phys. Rev. Phys. Educ. Res. 15, 020114 (2019).

[4] Adrian Madsen, Sarah B. McKagan, and Eleanor C. Sayre, Gender gap on concept inventories in physics: What is consistent, what is inconsistent, and what factors influence the gap?, Phys. Rev. ST Phys. Educ. Res. 9, 020121 (2013).

[5] Matthew Dew, Jonathan Perry, Lewis Ford, William Bassichis, and Tatiana Erukhimova, Gendered performance differences in introductory physics: A study from a large land-grant university, Phys. Rev. Phys. Educ. Res. 17, 010106 (2021).

[6] Brittany Bloodhart, Meena M. Balgopal, Anne Marie A. Casper, Laura B. Sample McMeeking, and Emily V. Fischer, Outperforming yet undervalued: Undergraduate women in STEM, PLoS One 15, e0234685 (2020).

[7] Linda J. Sax, Kathleen J. Lehman, Ramon S. Barthelem, and Gloria Lim, Women in physics: A comparison to science, technology, engineering, and math education over four decades, Phys. Rev. Phys. Educ. Res. 12, 020108 (2016).

[8] Emily M. Marshman, Z. Yasemin Kalender, Timothy Nokes-Malach, Christian Schunn, and Chandralekha Singh, Female students with A's have similar physics self-efficacy as male students with C's in introductory courses: A cause for alarm?, Phys. Rev. Phys. Educ. Res. 14, 020123 (2018).

[9] Z. Yasemin Kalender, Emily Marshman, Christian D. Schunn, Timothy J. Nokes-Malach, and Chandralekha Singh, Gendered patterns in the construction of physics identity from motivational factors, Phys. Rev. Phys. Educ. Res. 15, 020119 (2019).

[10] Z. Yasemin Kalender, Emily Marshman, Christian D. Schunn, Timothy J. Nokes-Malach, and Chandralekha Singh, Why female science, technology, engineering, and mathematics majors do not identify with physics: They do not think others see them that way, Phys. Rev. Phys. Educ. Res. 15, 020148 (2019).

[11] Z. Yasemin Kalender, Emily Marshman, Christian D. Schunn, Timothy J. Nokes-Malach, and Chandralekha Singh, Damage caused by women's lower self-efficacy on physics learning, Phys. Rev. Phys. Educ. Res. 16, 010118 (2020).

[12] Vashti Sawtelle, Eric Brewe, and Laird H. Kramer, Exploring the relationship between self-efficacy and retention in introductory physics, J. Res. Sci. Teach. 49, 1096 (2012).

[13] Katherine Rainey, Melissa Dancy, Roslyn Mickelson, Elizabeth Stearns, and Stephanie Moller, Race and gender differences in how sense of belonging influences decisions to major in STEM, Int. J. STEM Educ. 5, 10 (2018).

[14] Karyn L. Lewis, Jane G. Stout, Steven J. Pollock, Noah D. Finkelstein, and Tiffany A. Ito, Fitting in or opting out: A review of key social-psychological factors influencing a sense of belonging for women in physics, Phys. Rev. Phys. Educ. Res. 12, 020110 (2016).

[15] David Esparza, Amy E. Wagler, and Jeffrey T. Olimpo, Characterization of instructor and student behaviors in CURE and non-CURE learning environments: Impacts on student motivation, science identity development, and perceptions of the laboratory experience, CBE Life Sci. Educ. 19, ar10 (2020).

[16] Jared B. Stang and Ido Roll, Interactions between teaching assistants and students boost engagement in physics labs, Phys. Rev. ST Phys. Educ. Res. 10, 020117 (2014).

[17] Eric A. Williams, Justyna P. Zwolak, Remy Dou, and Eric Brewe, Linking engagement and performance: The social network analysis perspective, Phys. Rev. Phys. Educ. Res. 15, 020150 (2019).

[18] Adrienne L. Traxler, Tyme Suda, Eric Brewe, and Kelley Commeford, Network positions in active learning environments in physics, Phys. Rev. Phys. Educ. Res. 16, 020129 (2020).

[19] Justyna P. Zwolak, Remy Dou, Eric A. Williams, and Eric Brewe, Students' network integration as a predictor of persistence in introductory physics courses, Phys. Rev. Phys. Educ. Res. 13, 010113 (2017).

[20] Justyna P. Zwolak, Michael Zwolak, and Eric Brewe, Educational commitment and social networking: The power of informal networks, Phys. Rev. Phys. Educ. Res. 14, 010131 (2018).

[21] Cissy J. Ballen, Stepfanie M. Aguillon, Rebecca Brunelli, Abby Grace Drake, Deena Wassenberg, Stacey L. Weiss, Kelly R. Zamudio, and Sehoya Cotner, Do small classes in higher education reduce performance gaps in STEM? BioScience 68, 593 (2018).

[22] Cissy J. Ballen, Carl Wieman, Shima Salehi, Jeremy B. Searle, and Kelly R. Zamudio, Enhancing diversity in undergraduate science: Self-efficacy drives performance gains with active learning, CBE Life Sci. Educ. 16, ar56 (2017).

[23] D. Z. Grunspan, B. L. Wiggins, and S. M. Goodreau, Understanding classrooms through social network analysis: A primer for social network analysis in education research, Cell Biol. Educ. 13, 167 (2014).

[24] Angela M. Kelly, Social cognitive perspective of gender disparities in undergraduate physics, Phys. Rev. Phys. Educ. Res. 12, 020116 (2016).

[25] D. C. Haak, J. HilleRisLambers, E. Pitre, and S. Freeman, Increased structure and active learning reduce the achievement gap in introductory biology, Science 332, 1213 (2011).

[26] Ben Van Dusen and Jayson Nissen, Associations between learning assistants, passing introductory physics, and equity: A quantitative critical race theory investigation, Phys. Rev. Phys. Educ. Res. 16, 010117 (2020).

[27] Ben Van Dusen and Jayson Nissen, Equity in college physics student learning: A critical quantitative intersectionality investigation, J. Res. Sci. Teach. 57, 33 (2020).

[28] Jayson M. Nissen and Jonathan T. Shemwell, Gender, experience, and self-efficacy in introductory physics, Phys. Rev. Phys. Educ. Res. **12**, 020105 (2016).

[29] Stepfanie M. Aguillon, Gregor-Fausto Siegmund, Renee H. Petipas, Abby Grace Drake, Sehoya Cotner, and Cissy J. Ballen, Gender differences in student participation in an active-learning classroom, CBE Life Sci. Educ. **19**, ar12 (2020).

[30] Elise J. Dallimore, Julie H. Hertenstein, and Marjorie B. Platt, Leveling the playing field: How cold-calling affects class discussion gender equity, J. Educ. Learn. **8**, 14 (2019).

[31] Neil A. Lewis, Denise Sekaquaptewa, and Lorelle A. Meadows, Modeling gender counter-stereotypic group behavior: A brief video intervention reduces participation gender gaps on STEM teams, Social Psychol. Educ. **22**, 557 (2019).

[32] Jeffrey P. Adams, Gina Brissenden, Rebecca S. Lindell, Timothy F. Slater, and Joy Wallace, Observations of student behavior in collaborative learning groups, Astron. Educ. Rev. **1**, 25 (2001).

[33] Sarah L. Eddy, Sara E. Brownell, and Mary Pat Wenderoth, Gender gaps in achievement and participation in multiple introductory biology classrooms, CBE Life Sci. Educ. **13**, 478 (2014).

[34] Adrienne L. Traxler, Ximena C. Cid, Jennifer Blue, and Ramón Barthelemy, Enriching gender in physics education research: A binary past and a complex future, Phys. Rev. Phys. Educ. Res. **12**, 020114 (2016).

[35] Indigo Esmonde, Ideas and identities: Supporting equity in cooperative mathematics learning, Rev. Educ. Res. **79**, 1008 (2009).

[36] James A. Banks, Peter Cookson, Geneva Gay, Willis D. Hawley, Jacqueline Jordan Irvine, Sonia Nieto, Janet Ward Schofield, and Walter G. Stephan, Diversity within unity: Essential principles for teaching and learning in a multicultural society, Phi Delta Kappan **83**, 196 (2001).

[37] Ben Archibeque, Florian Genz, Maxwell Franklin, Scott V. Franklin, and Eleanor C. Sayre, Quantitative measures of equity in small groups, *Proceedings of the 2017 Physics Education Research Conference* (AIP, New York, 2018), pp. 44–47.

[38] Ben Archibeque, Mary-Bridget Kustusch, Florian Genz, Scott Franklin, and Eleanor C. Sayre, Qualitative measures of equity in small groups, in *Proceedings of the 13th International Conference of the Learning Sciences*, edited by Judy Kay and Rosemary Luckin (International Society of the Learning Sciences, London, United Kingdom, 2018), p. 1081.

[39] Idaykis Rodriguez, Eric Brewe, Vashti Sawtelle, and Laird H. Kramer, Impact of equity models and statistical measures on interpretations of educational reform, Phys. Rev. ST Phys. Educ. Res. **8**, 020103 (2012).

[40] Ben Van Dusen and Jayson Nissen, Equity in college physics student learning: A critical quantitative intersectionality investigation, J. Res. Sci. Teach. **57**, 33 (2020).

[41] Eric Burkholder, Cole Walsh, and N. G. Holmes, Examination of quantitative methods for analyzing data from concept inventories, Phys. Rev. Phys. Educ. Res. **16**, 010141 (2020).

[42] Heidi B. Carlone, Julie Haun-Frank, and Angela Webb, Assessing equity beyond knowledge- and skills-based outcomes: A comparative ethnography of two fourth-grade reform-based science classrooms, J. Res. Sci. Teach. **48**, 459 (2011).

[43] Joseph Kozminski, Heather Lewandowski, Nancy Beverly, Steve Lindaas, Duane Deardorff, Ann Reagan, Richard Dietz, Randy Tagg, M. EblenZayas, J. Williams *et al.*, *AAPT Recommendations for the Undergraduate Physics Laboratory Curriculum* (AAPT, College Park, MD, 2014).

[44] Paul W. Irving and Eleanor C. Sayre, Conditions for building a community of practice in an advanced physics laboratory, Phys. Rev. ST Phys. Educ. Res. **10**, 010109 (2014).

[45] Bethany R. Wilcox and H. J. Lewandowski, A summary of research-based assessment of students' beliefs about the nature of experimental physics, Am. J. Phys. **86**, 212 (2018).

[46] Emily M. Smith, Martin M. Stein, Cole Walsh, and N. G. Holmes, Direct Measurement of the Impact of Teaching Experimentation in Physics Labs, Phys. Rev. X **10**, 011029 (2020).

[47] Paul W. Irving and Eleanor C. Sayre, Identity statuses in upper-division physics students, Cult. Studies Sci. Educ. **11**, 1155 (2016).

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

[49] Danny Doucette, Russell Clark, and Chandralekha Singh, Hermione and the secretary: How gendered task division in introductory physics labs can disrupt equitable learning, Eur. J. Phys. **41**, 035702 (2020).

[50] James Day, Jared B. Stang, N. G. Holmes, Dhaneesh Kumar, and D. A. Bonn, Gender gaps and gendered action in a first-year physics laboratory, Phys. Rev. Phys. Educ. Res. **12**, 020104 (2016).

[51] 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).

[52] N. G. Holmes and Z. Yasemin Kalender, Preliminary evidence for available roles in mixed-gender and all-women lab groups, arXiv:2007.14833.

[53] Jasna Jovanovic and Sally Steinbach King, Boys and girls in the performance-based science classroom: Who's doing the performing? Am. Educ. Res. J. **35**, 477 (1998).

[54] Anna Teresia Danielsson and Cedric Linder, Learning in physics by doing laboratory work: Towards a new conceptual framework, Gender Educ. **21**, 129 (2009).

[55] Anna T. Danielsson, Exploring woman university physics students "doing gender" and "doing physics", Gender Educ. **24**, 25 (2012).

[56] Anna T. Danielsson, In the physics class: University physics students' enactment of class and gender in the context of laboratory work, Cult. Studies Sci. Educ. **9**, 477 (2014).

[57] Crystal L. Hoyt and Susan E. Murphy, Managing to clear the air: Stereotype threat, women, and leadership, Leadersh Q. **27,** 387 (2016).

[58] G. James Lemoine, Ishani Aggarwal, and Laurens Bujold Steed, When women emerge as leaders: Effects of extraversion and gender composition in groups, Leadersh Q. **27,** 470 (2016).

[59] L. S. Vygotsky, The dynamics of the schoolchild's mental development in relation to teaching and learning, J. Cogn. Educ. Psychol. **10,** 198 (2011).

[60] Michelene T. H. Chi and Ruth Wylie, The ICAP Framework: Linking cognitive engagement to active learning outcomes, Educ. Psychol. **49,** 219 (2014).

[61] Michelene T. H. Chi, Active-constructive-interactive: A conceptual framework for differentiating learning activities, Topics Cogn. Sci. **1,** 73 (2009).

[62] Michelene T. H. Chi, Joshua Adams, Emily B. Bogusch, Christiana Bruchok, Seokmin Kang, Matthew Lancaster, Roy Levy, Na Li, Katherine L. McEldoon, Glenda S. Stump, Ruth Wylie, Dongchen Xu, and David L. Yaghmourian, Translating the ICAP theory of cognitive engagement into practice, Cogn. Sci. **42,** 1777 (2018).

[63] M. K. Smith, W. B. Wood, W. K. Adams, C. Wieman, J. K. Knight, N. Guild, and T. T. Su, Why peer discussion improves student performance on in-class concept questions Science **323**, 122 (2009).

[64] Michelle K. Smith, Caleb Trujillo, and Tin Tin Su, The benefits of using clickers in small-enrollment seminar-style biology courses, CBE Life Sci. Educ. **10,** 14 (2011).

[65] Julie A. Bianchini, Where knowledge construction, equity, and context intersect: Student learning of science in small groups, J. Res. Sci. Teach. **34,** 1039 (1997).

[66] Sarah L. Eddy, Sara E. Brownell, Phonraphee Thummaphan, Ming-Chih Lan, and Mary Pat Wenderoth, Caution, student experience may vary: Social identities impact a student's experience in peer discussions. CBE Life Sci. Educ. **14,** ar45 (2015).

[67] Daniel Z. Grunspan, Sarah L. Eddy, Sara E. Brownell, Benjamin L. Wiggins, Alison J. Crowe, and Steven M. Goodreau, Males under-estimate academic performance of their female peers in undergraduate biology classrooms, PLoS One **11,** e0148405 (2016).

[68] Carl Wieman, Comparative cognitive task analyses of experimental science and instructional laboratory courses, Phys. Teach. **53,** 349 (2015).

[69] N. G. Holmes and Carl E. Wieman, Examining and contrasting the cognitive activities engaged in undergraduate research experiences and lab courses, Phys. Rev. Phys. Educ. Res. **12,** 020103 (2016).

[70] N. G. Holmes, Benjamin Keep, and Carl E. Wieman, Developing scientific decision making by structuring and supporting student agency, Phys. Rev. Phys. Educ. Res. **16,** 010109 (2020).

[71] Jennifer M. and Langer-Osuna, The social construction of authority among peers and its implications for collaborative mathematics problem solving, Math. Thinking Learning **18,** 107 (2016).

[72] Jennifer M. Langer-Osuna, How Brianna Became Bossy and Kofi came out smart: Understanding the trajectories of identity and engagement for two group leaders in a project-based mathematics classroom, Can. J. Sci. Math. Technol. Educ. **11,** 207 (2011).

[73] Sophia M. Jeon, Z. Yasemin Kalender, Eleanor C. Sayre, and N. G. Holmes, How do gender and inchargeness interact to affect equity in lab group interactions? in *Proceedings of the 2020 Physics Education Research Conference, virtual conference*, edited by Steven Wolf, Michael B. Bennett, and Brian W. Frank (AIP, New York, 2020), pp. 240–245.

[74] Sophia Jeon, Eleanor Sayre, and N. G. Holmes, What do equitable physics lab groups look like in light of inchargeness? in *Proceedings of the 14th International Conference of the Learning Sciences (ICLS) 2020*, Volume 2, edited by Melissa Gresalfi and Ilana Seidel Horn (International Society of the Learning Sciences, Nashville, TN, 2020), pp. 777–778.

[75] Mary Bridget Kustusch, Eleanor C. Sayre, and Scott Franklin, Identifying shifts in agency by analyzing authority in discussion, in *Proceedings of the 13th International Conference of the Learning Sciences*, edited by Judy Kay and Rosemary Luckin (International Society of the Learning Sciences, London, United Kingdom, 2018), p. 1623.

[76] Patricia Heller and Mark Hollabaugh, Teaching problem solving through cooperative grouping. Part 2: Designing problems and structuring groups, Am. J. Phys. **60,** 637 (1992).

[77] K. D. Tanner, Structure matters: Twenty-one teaching strategies to promote student engagement and cultivate classroom equity, Cell Biol. Educ. **12,** 322 (2013).

[78] N. Ding and E. Harskamp, How partner gender influences female students' problem solving in physics education, J. Sci. Educ. Technol. **15,** 331 (2006).

[79] https://diversity.cornell.edu/reports-and-data

[80] *Data on People from Underrepresented Groups in the Physical Sciences and Engineering*, Technical Report (American Institute of Physics, New York, 2020), https://www.aip.org/statistics/resources/data-underrepresented-groups-physical-sciences-and-engineering.

[81] Patrick J. Mulvey and Starr Nicholson, *Physics Bachelor's Degrees: 2018*, Technical Report (American Institute of Physics, New York, 2020).

[82] Simone Hyater-Adams, Claudia Fracchiolla, Noah Finkelstein, and Kathleen Hinko, Critical look at physics identity: An operationalized framework for examining race and physics identity, Phys. Rev. Phys. Educ. Res. **14,** 010132 (2018).

[83] Simone Hyater-Adams, Claudia Fracchiolla, Tamia Williams, Noah Finkelstein, and Kathleen Hinko, Deconstructing Black physics identity: Linking individual and social constructs using the critical physics identity framework, Phys. Rev. Phys. Educ. Res. **15,** 020115 (2019).

[84] Katemari Rosa and Felicia Moore Mensah, Educational pathways of Black women physicists: Stories of experiencing and overcoming obstacles in life, Phys. Rev. Phys. Educ. Res. **12,** 020113 (2016).

[85] Angela Johnson, Jaweer Brown, Heidi Carlone, and Azita K. Cuevas, Authoring identity amidst the treacherous terrain of science: A multiracial feminist examination of

the journeys of three women of color in science, J. Res. Sci. Teach. **48**, 339 (2011).

[86] Rachel E. Scherr, Mike A. Lopez, and Marialis Rosario-Franco, Isolation and connectedness among Black and Latinx physics graduate students, Phys. Rev. Phys. Educ. Res. **16**, 020132 (2020).

[87] Saeed Moshfeghyeganeh and Zahra Hazari, Effect of culture on women physicists' career choice: A comparison of Muslim majority countries and the West, Phys. Rev. Phys. Educ. Res. **17**, 010114 (2021).

[88] Zahra Hazari, Philip M. Sadler, and Gerhard Sonnert, The science identity of college students: Exploring the intersection of gender, race, and ethnicity, J. Coll. Sci. Teach. **42**, 82 (2013), https://www.jstor.org/stable/43631586.

[89] Sreyashi Jhumki Basu and Angela Calabrese Barton, Critical physics agency: Further unraveling the intersections of subject matter knowledge, learning, and taking action, Cult. Studies Sci. Educ. **4**, 387 (2009).

[90] Stephen Kanim and Ximena C. Cid, Demographics of physics education research, Phys. Rev. Phys. Educ. Res. **16**, 020106 (2020).

[91] Allison J. Gonsalves, Anna Teresia Danielsson, and Helena Pattersson, Masculinities and experimental practices in physics: The view from three case studies, Phys. Rev. Phys. Educ. Res. **12**, 020120 (2016).

[92] N. G. Holmes, Carl E. Wieman, and D. A. Bonn, Teaching critical thinking, Proc. Natl. Acad. Sci. U.S.A. **112**, 11199 (2015).

[93] N. G. Holmes and Emily M. Smith, Operationalizing the AAPT learning goals for the lab, Phys. Teach. **57**, 296 (2019).

[94] Z. Y. Kalender, E. Stump, K. Hubenig, and N. G. Holmes, Restructuring physics labs to cultivate sense of student agency, Phys. Rev. Phys. Educ. Res. **17**, 020128 (2021).

[95] Sapna Cheryan, Sianna A. Ziegler, Amanda K. Montoya, and Lily Jiang, Why are some STEM fields more gender balanced than others?, Psychol. Bull. **143**, 1 (2017).

[96] Sarah Banchefsky, Jacob Westfall, Bernadette Park, and Charles M. Judd, But you don't look like a scientist!: Women scientists with feminine appearance are deemed less likely to be scientists, Sex Roles **75**, 95 (2016).

[97] Brian A Nosek *et al.*, National differences in gender-science stereotypes predict national sex differences in science and math achievement, Proc. Natl. Acad. Sci. U.S.A. **106**, 10593 (2009).

[98] Brian A. Nosek and Frederick L. Smyth, Implicit social cognitions predict sex differences in math engagement and achievement, Am. Educ. Res. J. **48**, 1125 (2011).

[99] Asia A. Eaton, Jessica F. Saunders, Ryan K. Jacobson, and Keon West, How gender and race stereotypes impact the advancement of scholars in STEM: Professors' biased evaluations of physics and biology post-doctoral candidates, Sex Roles **82**, 127 (2020).

[100] Sarah Webb, The importance of undergraduate research, Science (2007).

[101] Yang Yang, Nitesh V. Chawla, and Brian Uzzi, A network's gender composition and communication pattern predict women's leadership success, Proc. Natl. Acad. Sci. U.S.A. **116**, 2033 (2019).

[102] Amy D. Robertson, S. B. McKagan, and Rachel E. Scherr, Selection, generalization, and theories of cause in physics education research: Connecting paradigms and practices, in *Getting Started in PER*, edited by Charles R. Henderson, and Kathleen A. Harper (American Association of Physics Teachers, College Park, MD, 2018).

[103] Patricia Heller, Thomas Foster, and Kenneth Heller, Cooperative group problem solving laboratories for introductory classes, AIP Conf. Proc. **399**, 913 (1997).

[104] Kimberly Tanner, Liesl S. Chatman, and Deborah Allen, Approaches to cell biology teaching: Cooperative learning in the scienceclassroom—Beyond students working in groups, Cell Biol. Educ. **2**, 1 (2003).

[105] Elli J. Theobald *et al.*, Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math, Proc. Natl. Acad. Sci. U.S.A. **117**, 6476 (2020).

[106] Sat Gavassa, Rocio Benabentos, Marcy Kravec, Timothy Collins, and Sarah Eddy, Closing the achievement gap in a large introductory course by balancing reduced in-person contact with increased course structure, CBE Life Sci. Educ. **18**, ar8 (2019).

[107] Connor Neill, Sehoya Cotner, Michelle Driessen, and Cissy J. Ballen, Structured learning environments are required to promote equitable participation, Chem. Educ. Res. Pract. **20**, 197 (2019).

[108] Zahra Hazari, Cheryl Cass, and Carrie Beattie, Obscuring power structures in the physics classroom: Linking teacher positioning, student engagement, and physics identity development, J. Res. Sci. Teach. **52**, 735 (2015).

[109] Peter Wulff, Zahra Hazari, Stefan Petersen, and Knut Neumann, Engaging young women in physics: An intervention to support young women's physics identity development, Phys. Rev. Phys. Educ. Res. **14**, 020113 (2018).