JMBE Journal of Microbiology & Biology Education

Participation and Performance by Gender in Synchronous Online Lectures: Three Unique Case Studies during Emergency Remote Teaching

Sierra C. Nichols,[a] Yongyong Y. Xia,[a] Mikaylie Parco,[a] and Elizabeth G. Bailey[a]
[a]Department of Biology, Brigham Young University, Provo, Utah, USA

Studies have documented that men's voices are generally heard more than women's voices in face-to-face undergraduate biology classes, and some performance gaps have also been documented. Some of the few studies on gender equity in traditional online biology education suggest that women participate more and perform better in asynchronous online courses compared to men. While much is known about emergency remote teaching during the COVID-19 pandemic exacerbating existing inequities generally, studies are needed about the impact of specific emergency remote teaching practices on specific groups such as women. In this study, we performed an in-depth investigation of three life sciences classrooms that utilized synchronous online lectures during the pandemic. We observed each class throughout the semester, quantified participation behaviors, and investigated the role of student gender. We also compared final course grades by gender. On average, we found that men participated more than women both verbally and by chat. These differences were not significant for each class individually, but the differences align with the face-to-face patterns seen in this population previously. Our results also hint that men's chat comments may be more likely to be acknowledged than women's chats by peers. We found evidence of greater performance disparities favoring men than seen previously before the pandemic, but not in all classes. We discuss implications for instructors conducting emergency remote teaching as well as the need for larger studies to test the replicability of our results.

KEYWORDS gender gaps, achievement, verbal participation, Zoom, videoconference, women in STEM, female, undergraduate, biology education, COVID

INTRODUCTION

Women are underrepresented in science, technology, engineering, and math (STEM), but these fields have become more equitable over time. In the last 50 years, the number of women in the workforce in STEM fields increased from 8% to 27% (1). Women are well represented in biology, receiving more than half of the degrees awarded in 2016 at the bachelor's, master's, and doctorate levels (2). However, a closer look at the natural sciences reveals subtle differences referred to as the leaky pipeline, as we see fewer women represented along the path toward tenured professorship (3–6). At the same time, demand for educated and diverse workers is continually increasing, and increased inclusion of women could help meet that demand (7). Research suggests that gender diversity benefits companies and organizations, because women's unique experiences are likely to give them different perspectives and ideas (8, 9). Thus, it is important that we strive to make STEM education more equitable and close opportunity gaps for women (7).

As discussed by Eddy and Brownell (10), equal enrollment numbers do not guarantee gender parity, and the effects of undergraduate student performance, engagement, and student affect have important implications down the pipeline. For example, women may recognize gender inequities in classroom participation more than men and thus note classroom environments that are more unwelcoming (11). These classroom experiences can have long-lasting effects, as women are less visible and lack social standing when they do not speak up (12) and students' sense of belonging in STEM impacts their STEM career interests (13). Furthermore, participation can impact performance, as students' perceptions of engagement, including active or collaborative learning and faculty interaction, are strongly related to their grades (14). Again, classroom performance can have long-term effects as grades and perceptions of ability or performance are strong predictors of persistence (15, 16). Thus, we chose to focus on undergraduate women's in-class participation and course performance.

Participation and performance gender inequities in undergraduate biology classrooms

Recent studies on traditional in-person biology classrooms have documented that men are more likely to speak up than

Address correspondence to Department of Biology, Brigham Young University, Provo, Utah, USA. E-mail: liz_bailey@byu.edu.
Sierra C. Nichols, Yongyong Y. Xia, and Mikaylie Parco are undergraduate student authors.
The authors declare no conflict of interest.
Received: 1 October 2021, Accepted: 16 November 2021, Published: 28 March 2022

women in class (17–21). Some studies have found performance gender gaps in favor of men in introductory biology classes, especially on exams requiring higher-level cognitive skills (18, 22), while others found either no difference by gender (23) or mixed results (24). We previously found that gender gaps in performance varied across a wide range of life sciences courses, and that women were more likely to underperform compared to men in courses instructed by a man and/or where women students were the minority (17). A recent meta-analysis found that large class sizes, high reliance on exam scores, and lack of active learning strategies were all predictive of decreased women's performance compared to men in undergraduate biology and chemistry classes (25).

Emergency remote teaching during the covid-19 pandemic

During the global COVID-19 pandemic, universities canceled face-to-face classes and transitioned to online learning. This instruction has been termed emergency remote teaching (ERT) to contrast this sudden and temporary change from traditional online learning (26). Students have reported some benefits associated with ERT such as flexibility (27, 28), but the literature also lists many challenges (27–31). We note that difficulties arising from the pandemic were not felt by all populations equally, and existing inequalities were exacerbated (31–34). Thus, we must investigate how ERT affected marginalized underrepresented groups in STEM, including women.

Gender inequities in ERT and traditional online learning

Recent research found that women were more likely to report decreased motivation, focus, and self-direction than their male peers during ERT (31). Furthermore, studies suggest that women are more likely to turn off their camera due to concerns about their background (35), and women may get zoom fatigue due to mirror anxiety (i.e., anxiety triggered by the perpetual mirror during video conferences) more than men (36). Reinholz et al. (37) studied biology classrooms that started in person and then transitioned online mid-semester and did not find evidence of widened gender inequities. They suggest that the diverse Zoom features provided more participation spaces for women.

Much of the previous research on traditional online learning has been done on courses that utilize asynchronous activity, and most results suggest that women have a more positive perception of online classes than males, especially in terms of communication and connectedness, and women often perform better than men online (38–40). In a large study that compared students' online performance to previous in-person experiences, researchers found that students generally performed worse in online settings, especially men (41). Many women described anonymity and asynchronous aspects of the course as the most positive aspects (42). Many studies on asynchronous online discussions report that women post more than men and show higher engagement and collaboration (43, 44); however, some studies have found no gender differences in post rates (45).

Our research questions

While there is a growing body of research on gender inequities in performance and in-class participation for biology classes, the impact of ERT practices on gender equities in undergraduate biology classes is largely unknown. We chose to investigate synchronous online lectures, as these are the most like face-to-face classroom experiences, and performed an in-depth analysis of three life sciences classrooms to investigate these research questions:

1. How are instructors inviting participation in synchronous online lectures during ERT?
2. Are men and women equally likely to participate in synchronous online lectures overall? Verbally? By chat? Does this depend on the gender ratios in the class?
3. Do these participation trends change over the course of the semester?
4. Are there gender differences in final course grades earned in ERT contexts that use synchronous lectures?

METHODS

Ethics statement and participant inclusion criteria

This study was approved by the Brigham Young University Institutional Review Board (Protocol E2020-356). After receiving approval, we emailed the instructors of 61 synchronous remote classes in the life sciences during the 2020–2021 school year when ERT was common. Fourteen of those instructors agreed to let us seek consent from their students to consider their class, and all students consented for Zoom recordings to be released in six of those classes. Three of those six classes were excluded because there were fewer than three men enrolled. The final inclusion criteria were that the course must be an undergraduate class in the life sciences, utilize live/synchronous online lectures, have at least three men enrolled, and the instructor and all students needed to consent for videos to be viewed. While we originally intended to replicate the scale of our previous study on in-person classrooms (17) ($n = 34$), we ended up with a smaller number of classes ($n = 3$). This allowed us to do a much more in-depth study of these classes.

Study population

Our study took place at a private, religious R2 university. The student population is broadly Christian, relatively conservative, and generally high performing due to the competitive nature of enrollment. In the college, ∼50% of the students are women and ∼15% of the faculty are women (internal data, 2019). The course characteristics for the three classes used in this study are summarized in Table 1, and more details about

TABLE 1
Course characteristics

Characteristic	Class 1	Class 2	Class 3
Class level	400-level	400-level	100-level
Audience	Majors	Majors	General Education
Instructor gender	Female	Male	Male
Teaching assistent genders	Female	Female	Female
Class size	47 students	22 students	67 students
Female	25 (53%)	13 (59%)	48 (72%)
Male	22 (47%)	9 (41%)	19 (28%)
Chat/verbal participation required?	No	Yes	No
Percent of final grade determined subjectively	~33%	~66%	~70%
Percent of final grade dependent on exam scores	50%	43%	30%

each course are found in the Supplemental Materials. In Table 1, "points awarded subjectively" refers to points earned on open response questions, essays, etc. as opposed to multiple choice or similar question types that are computer graded. These subjective points could be on exams or lower-stakes homework assignments.

Classroom observation methods

We observed class sessions by analyzing video recordings and chat box text files from Zoom videoconferencing software. We collected recordings of three lectures that took place early in the semester (first 5 weeks of a 15-week semester), three from mid-semester, and three from late in the semester (last 3 weeks of semester). For Class 1, we were unable to get recordings early in the semester, so only mid-semester and late semester data are included. For some analyses, results from all six (Class 1) or nine (Class 2 and Class 3) observations were averaged or combined to get an overall look at the whole semester. In other analyses, results from three observations were averaged or combined to give an average look at participation at a specific point in the semester (early, mid, or late).

COPUS

First, we used the Classroom Observations Protocol for Undergraduate STEM (COPUS) to analyze each recording (46, 47). Twenty-five unique codes were used to quantify instructor and student behaviors in 2-min intervals. We also collapsed those 25 codes into four codes for instructors and four codes for students (48) (see Appendix 1).

Two researchers (S.C.N. and Y.Y.X.) watched and coded the first video for each class individually, and we assessed inter-rater reliability. Some of the COPUS codes were never coded by our researchers (e.g., students presenting to the entire class). Cohen's kappa for the rest of the 25 codes ranged from 0.835 to 1.000. Cohen's kappa for the eight collapsed codes ranged from 0.913 to 1.000. As raters had excellent agreement, each subsequent video was then only coded by one researcher.

We report COPUS results in a primarily descriptive manner with collapsed codes (48), as we did not have a large data set on which we could perform cluster analysis as has been suggested (46). We also looked at specific codes, including percent of time spent lecturing, to compare our data to previous cluster analyses which grouped courses into didactic, interactive lecture, and student-centered classrooms (49). However, we recognize the limitations of using the cluster analysis from a separate data set and merely use these categories to discuss the differences between our three classes.

Instructor invitations and student participation

Next, we quantified student participation and instructor invitations to participate based on a protocol adapted from our previous in-person classroom observations (17). For each video, we recorded the number of times the instructor made an explicit invitation for verbal participation, invited students to use the chat feature, put students in breakout rooms, and polled the class. More detailed definitions of these behaviors are found in Appendix 1.

We also recorded the number of times each student typed in the chat box or participated verbally. We also noted whether students' chat comments were acknowledged (either verbally or in another chat comment) by instructors (the course instructor or TAs) or peers. In the recordings, we could not see a participant list, so we could not tell whether students on the roster were logged into Zoom or not. Thus, we did not have a clear way to track absence vs silence, so non-participators in our study may either be absent students or attending students who never participated. More detailed definitions of these behaviors are found in Appendix 1.

We assumed and recorded each student's gender based solely on whether they were male- or female- presenting in photos or their voice. This is clearly a limitation of our study, as we do not know the possible complexities of each student's gender

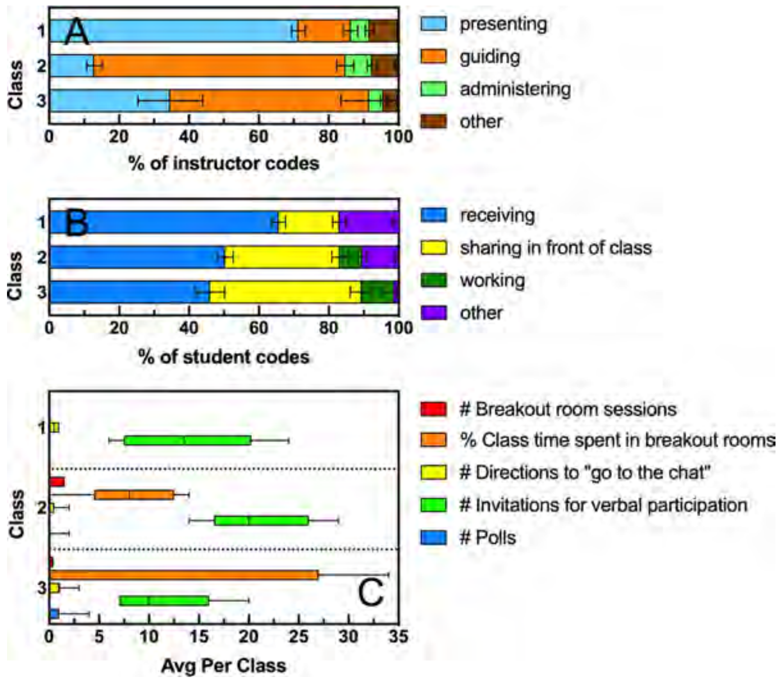

FIG 1. Teaching style based on instructor and student behaviors. Panels A and B: Zoom recordings of synchronous lectures were coded using the Classroom Observation Protocol for Undergraduate STEM (COPUS). All codes were collapsed into four categories for instructor behaviors (Panel A) and four categories for student behaviors (Panel B). The *x* axes represent the percent of all codes. Data are averaged from six (Class 1) or nine (Classes 2 and 3) observations, and error bars represent standard error of the mean. Panel C: Instructor behaviors that invite student participation were quantified for each video recording ($n = 6$ for Class 1 and $n = 9$ for Class 2 and Class 3).

identity. The numbers of men and women in each class determined in this way matched the numbers of males and females in the lists obtained from the registrar's office, which include the sex listed on students' educational records (see Table 1).

When recording settings did not save student names, we matched verbal participators to the pictures provided on the course picture roll. If the face was not an obvious match, we consulted with the research team to come to a consensus on which student was participating. Because we had six or nine observations of each class, students who participated with their cameras off occasionally were able to be identified based on voice comparisons with other observations in which their camera was turned on.

Three raters (Y.Y.X., S.C.N., M.P.) analyzed the first video for each class. Initially, interrater reliability (Cohen's kappa) ranged from 0.347 to 0.762. Raters then met and talked through differences until they came to a consensus. The three raters then all rated the same video and achieved near perfect alignment (Y.Y.X. and S.C.N. kappa = 0.98; Y.Y.X. and MP kappa = 0.96; S.C.N. and M.P. kappa = 0.98). Most of the rest of the videos were then coded by a single rater. Occasionally a video had so much participation that two raters would work together to properly identify all students.

Course grades

We obtained final letter grades for all students in each of the three courses from the university registrar's office. Data were de-identified except for gender (biological sex on educational records) and ACT score. Thus, we can compare final course performance by gender overall for each course, but we cannot match specific students' participation behavior with their course performance. Final letter grades (e.g., A, A–, B+, etc.) were converted to a 4-point scale (e.g., A = 4, A– = 3.7, B+ = 3.3, etc.). When grades are presented as ACT-adjusted grades, these are the estimated marginal means for gender after performing a one-way ANCOVA. The one-way ANCOVA was performed for each class separately, with gender as the independent variables and ACT score as the covariate as an imperfect measure of preparation.

RESULTS

Course characteristics

First, we analyzed video recordings of synchronous zoom lectures using COPUS (see Methods). As shown in Fig. 1A and B, the three courses we observed were all unique in terms of teaching style. Class 1 was more didactic with ~70% of instructor codes fitting in the "presenting" category (Fig. 1A; blue). Class 2 was more student-centered with ~70% of instructor codes fitting in the "guiding" category (Fig. 1A; orange). Finally, Class 3 was more of an interactive lecture, but there was more variety from class period to class period. On average, student codes were evenly split between "receiving" and "sharing in front of class" (Fig. 1B; blue and yellow, respectively).

We quantified instructor behaviors that encouraged participation for each observation. As shown in Fig. 1C, instructors of all three courses used a lot of invitations for students to unmute and participate verbally (green), but they rarely explicitly directed the students to use the chat feature. Only the instructors of Class 2 and Class 3 utilized Zoom breakout rooms (red = # sessions; orange = % class time spent in breakout rooms). None of the instructors used Zoom polls frequently (blue), and we noted that the instructors of Class 2 and Class 3 only used these in the observations that took place early in the semester.

Student performance

We acquired final course grades for each class from the registrar's office to compare student performance by gender. In terms of raw grades, male students significantly outperformed their female peers in Class 1 (Fig. 2A $P = 0.04$ by unequal variance *t* test), and this difference of about a half of a letter grade was virtually unchanged after adjusting for ACT score (Fig. 2B $P = 0.05$ by *unpaired t test*). In Class 2, males also outperformed females in the course by about a full letter grade (Fig. 2C), but the difference was only significant for ACT-adjusted scores (Fig. 2D $P = 0.007$ by *unpaired t test*). Finally, males and females earned equivalent grades in Class 3 (Fig. 2E and F).

Student participation

Next, we quantified student participation from recordings of synchronous Zoom lectures. As shown in Fig. 3, overall participation and gender comparisons varied by class. In Class 1, about 60% of the students never participated, but this was true for both men and women (Fig. 3A). Those that did participate were equally likely to choose chat or verbal participation regardless of gender (Fig. 3A). Furthermore, we saw no difference in the average number of chat comments (Fig. 3B) or verbal participation events (Fig. 3C) by gender in Class 1, although the students who participated the most were men (Fig. 3B and C). All participation in Class 2 was verbal for almost all men and women, and the only two students who never participated were women (Fig. 3D). Both chat and verbal event rates were indistinguishable by gender (Fig. 3E and F), but chatting was very rare in this class. Class 3 utilized both the chat feature and verbal participation (Fig. 3G), and women may have been more likely to choose chat over verbal participation and more likely to not participate at all, $\chi^2(3) = 5.12$ ($P = 0.16$). However, based on average number of events per observation, men and women were equally likely to chat (Fig. 3H) and participate verbally (Fig. 3I). Interestingly, the students who participated the most in this class were women (Fig. 3I). For comparison to other studies (17, 20), the rate ratios (average women's rate/average men's rate) are 0.30, 0.70, 0.12, 0.63, 0.78, and 0.91 (Fig. 3B, C, E, F, H, and I, respectively).

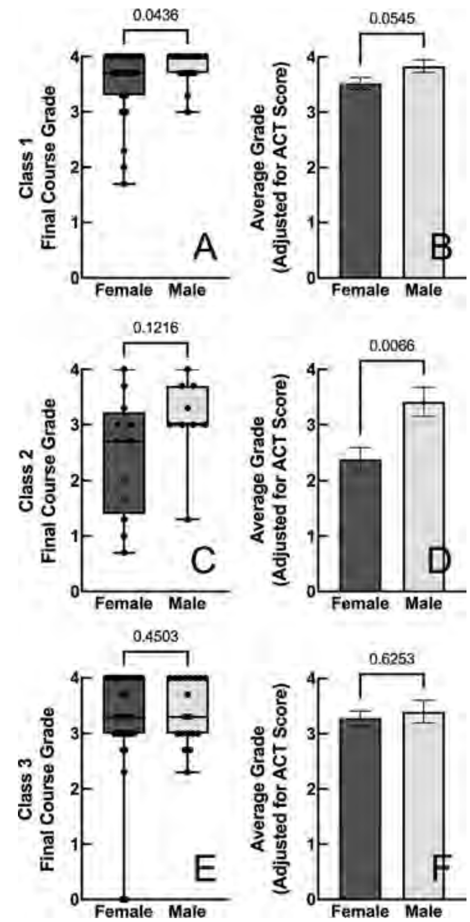

FIG 2. Final course grades by gender. Final course grades for Class 1 (Panels A and B), Class 2 (Panels C and D), and Class 3 (Panels E and F) were obtained from the university registrar and were converted to a four-point scale. Panels A, C, and E show raw grade distributions (minimum to maximum with quartiles) by gender. Panels B, D and F show estimated marginal means after adjusting for ACT score (mean ± standard error). Brackets show *P* value after *unpaired t test* for each panel, with Panels A and E requiring an unequal variances *t* test.

We were also able to track the participation of each student over time, and these detailed data are shown in Appendix 2 (Fig. S1–S3). These heat maps show the complex landscape of class participation as a function of time. Class 1 participation is dominated by a few students, especially one man (Fig. S1). Class 2 has much more spread-out participation among most students, although one man and one woman still dominate the discussion (Fig. S2). Class 3 is somewhat in the middle with many silent students (mostly women), and a few women who were the highest participators (Fig. S3).

We noticed that in Fig. S1, only two women participated in any of the three mid-semester observations, but many women participated later in the semester. We quantified the proportion of men and women who participated at least once at different phases in the semester, and the data are shown in Fig. 4. By Fisher's exact test, women were significantly less likely to participate than men mid-semester ($P = 0.03$) but perhaps more likely to participate than men

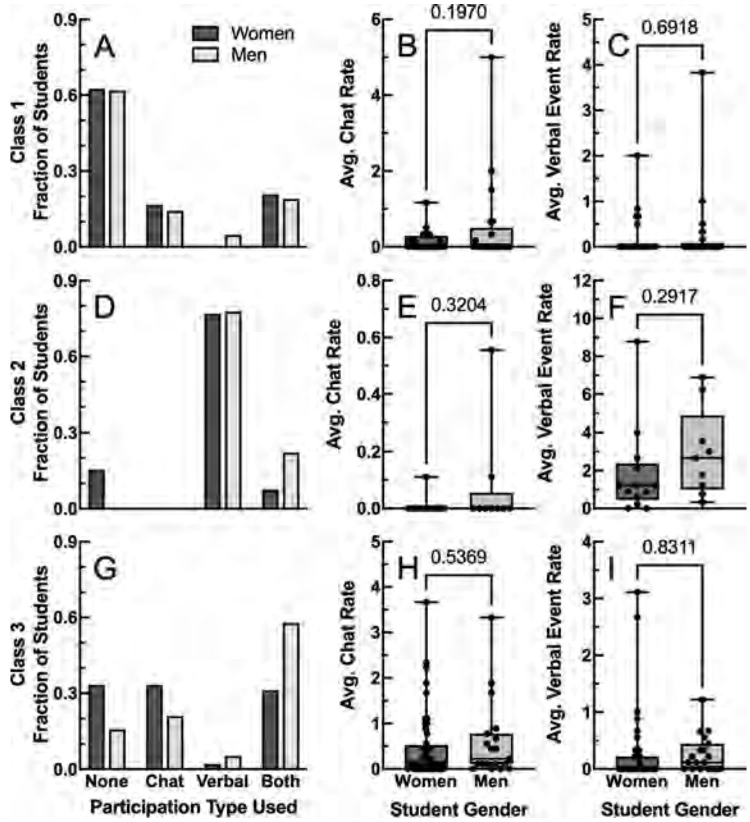

FIG 3. Classroom participation by gender. All instances of student participation (typed chats or verbal) were recorded from videos of synchronous Zoom lectures for Class 1 (Panels A–C), Class 2 (Panels D–F), and Class 3 (Panels G–I). Panels A, D, and G: Students were classified as participating both verbally and by chat at least once each over all observations ("Both"), at least once verbally but never by chat ("Verbal"), at least once by chat but never verbally ("Chat"), or never participating in any observation ("None"). Fraction of students in each category is shown on the *y* axis by gender. Panels B, E, and H: For each student, chat rate was calculated as the average number of chat comments per observation. Panels C, F, and I: For each student, verbal event rate was calculated as the average number of verbal comments per observation. Rates are compared by gender, and brackets show *P* value after unequal variances *t* test (Panels B, C, E, and I) or unpaired *t* test (Panels H and F).

later ($P = 0.20$). Our results also would have looked different had we only looked at a specific point in the semester for the other classes. In Class 2, women participated less and less as the semester went on (Fig. S2), but this could also be true for men. In Class 3, women may have been least represented in classroom participation in the middle of the semester (Fig. S3).

Finally, we investigated whether men's and women's chat comments were equally likely to be acknowledged by instructors and peers. Due to the limited use of the chat feature in Classes 1 and 2 (see Fig. 3), we focused on Class 3. As shown in Fig. 5A, chats of men and women were equally likely to be acknowledged by instructors. However, we saw a hint that men's chats may get acknowledged more frequently by peers than women's chat comments (Fig. 5B), with a few men's chats getting acknowledged more frequently than anyone else's. We saw no association between acknowledgments and future participation rates (data not shown).

DISCUSSION

We did an in-depth analysis of ERT in three life sciences classrooms that had synchronous lectures. These three courses were very different in context, composition, and instructional style (Table 1, Fig. 1). We can compare results from this case study of ERT to those from in-person courses published on our population previously (17).

Student performance

With women making up about 70% of students enrolled in male-instructed Class 3, the lack of a gender difference in final course grades (Fig. 2E and F) aligns with what we would predict for in-person courses in our population (see Fig. 6 in Ref. 17). However, we saw much larger gender gaps in final course grades for Class 1 (about a half letter grade in favor of males; Fig. 2A and B) and Class 2 (about a full letter

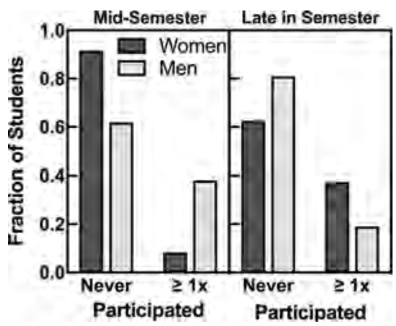

FIG 4. Class 1 participation representation by gender changed during semester. The fraction of students who participated at least once (including chats and verbal participation) was calculated based on Zoom lecture video observations. Mid-Semester data came from the three observations in the middle of the semester, and Late in Semester data came from the three observations near the end of the semester. Students had to participate at least once in any of the three lectures to be counted in the $\geqslant$1x category. Men were more likely to participate mid-semester (Fisher's Exact Test: $P = 0.03$, $n = 45$), but women participated more later in the semester (Fisher's Exact Test: $P = 0.20$, $n = 45$).

grade in favor of males, Fig. 2C and D). From our previous data set (see Fig. 6 in Ref. 17), we would have predicted females to earn higher grades than males in a female-instructed class with about 50% female students like Class 1, and we would have predicted no gender gap in a male-instructed class with about 60% female students like Class 2. This suggests that ERT classrooms may exacerbate gender achievement gaps. Class 1 results would align with our previous results if it had been taught by a man, so perhaps the benefits of female role models are reduced online without in-person relationships. Our results do align with a recent meta-analysis (25) in that Classes 1 and 2 had greater dependence on exam scores (Table 1) and underperformance of women (Fig. 2A–D). Our results differ from some studies in traditional online learning that found women perform better than men (40).

Student participation

Unsurprisingly, more active classrooms (Classes 2 and 3; Fig. 1) provided opportunities for a larger proportion of all students to be involved and participate in class (Fig. 3D, G, Fig. S2 and S3). Our most didactic class (Class 1; Fig. 1) had a much smaller proportion of students participating (Fig. 3A and Fig. S1). We saw no clear association between how active a classroom was and gender equity in participation, as women may have been more likely to stay silent than men in Class 3 (Fig. 3G), but no clear gender pattern was seen in Class 1 (Fig. 3A). This could possibly be explained by the slightly larger class size in Class 3 (Table 1); (20), but Class 3 also had more women, which has been shown to predict more gender equity (17). Perhaps having more female peers adds belonging in person when these peers are consistently visible, but invisible female peers in an online lecture may not have the same effect. The degree to which female peers are "invisible" would depend on camera use in the class and whether students were visible to each other. Unfortunately, we could not quantify the fraction of cameras that were on because Zoom recordings did not show the full gallery view when the instructor was sharing their screen (which applied most of the time).

In terms of in-class participation rates, no class showed statistical differences between men and women (Fig. 3); however, in our previous study on this population, small participation gaps were statistically reproducible over many classes and predicted by student gender ratios (17). If we compare our rate ratios (women's average rate/men's average rate) with those in that original study (see Fig. 4A in Ref. 17) for classes with 50% females (Class 1), 60% females (Class 2), and 70% females (Class 3), our verbal rate ratios of 0.7, 0.63, and 0.91, respectively, align well with those original results. Thus, we did not find any evidence that gender gaps in verbal participation are any different for synchronous online lectures versus in-person lectures. Interestingly, the chat rate ratios (women/men) were much lower than our in-person verbal data from 2020 for Class 1 (0.30) and Class 2 (0.12), suggesting there could be larger gender gaps in favor of men for typed chats. However, using the chat feature was not the norm in these classes, and chat events were very rare. In Class 3, where chatting was the norm, the chat rate ratio (women/men) of 0.78 aligns well with our previous data. Thus, we believe that the chat feature is as equitable as verbal participation if it is commonly used according to class norms.

If we look at the fraction of students who used the chat and/or unmuted in Class 3 (where both methods were common; Fig. 3G), we see hints that women may choose the chat as opposed to participating verbally. As proposed previously (37), using a variety of different Zoom features could possibly facilitate a more friendly and gender equitable environment. Women have been shown to use asynchronous communication channels more than men in traditional online learning to

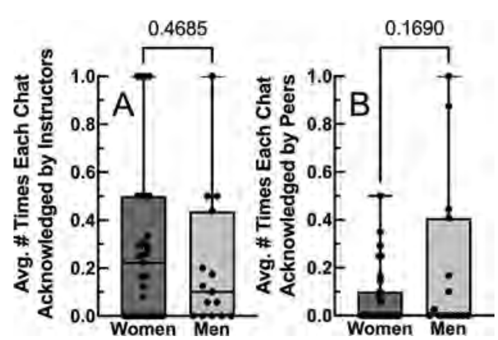

FIG 5. Average number of times students' chats were acknowledged by instructors or peers in Class 3. For all students who chatted at least once during the nine observations of Class 3 (31 women and 15 men), we calculated the average number of times their chat was acknowledged by instructors (Panel A; includes both the primary instructor and the TAs) or peers (Panel B). Distributions (minimum to maximum with quartiles) are shown by gender. Brackets show P value after *unpaired t test* (Panel A) or unequal variances t test (Panel B).

create online community (43, 44), so it is possible women are doing the same thing with the chat feature.

In our previous study (17), participation and performance gender gaps were not associated with each other. Here, we saw a possible correlation between participation gender equity (Fig. 3) and performance gender equity (Fig. 2), as Class 3 had the least difference between grades by gender and the participation rate ratio closest to 1. As discussed above, there are other factors that we would expect to affect grade differences by gender, but participation trends in class might also contribute.

Our study design allowed us to look systematically at different points in the semester and observe each class six to nine times. Previous studies have observed classes anywhere from two to 20 times, but most classes were observed closer to three times each (17–20). We found that conclusions on gender equity in participation may differ based on how many observations you make and at what point in semester you make them, as we would have drawn different conclusions with fewer observations at specific points in semester (Fig. 4, Fig. S1–S3).

Finally, we saw hints that men's chat comments may be acknowledged by peers more than those made by women (Fig. 5B). This aligns with research that suggests men are likely to overestimate academic ability of male peers and underestimate ability of female peers (50). Male and female faculty can also have biases that favor men (51). However, we have not seen evidence of faculty explicitly favoring men's participation in the past (17), nor did we see instructors acknowledging men's or women's chats differently here (Fig. 5A).

Limitations

First, we acknowledge that we were unable to have students self-disclose their gender identity. Because most students did not have their pronouns labeled on their display name, we had to rely on stereotypical features to classify students as male- or female-presenting. Thus, we do not know the complexities with which students identify based on gender. Furthermore, we obtained de-identified grades from the registrar's office which were categorized according to biological sex, which we acknowledge may or may not match the gender which students present. Our conclusions are limited by our inability to differentiate between absent students and attending students who remain silent. As ERT may lead to less consistent class attendance, gender differences in attendance habits could confound our results. Our small sample size is also a limitation. Future replication studies should help identify which of our results are generalizable patterns.

Implications for instructors

Based on our results, we encourage instructors to have more interactive and student-centered classrooms. Utilizing a greater variety of Zoom features (e.g., breakout rooms, polls, and the chat feature) will likely increase the proportion of students who get to participate, and women may be more likely to participate via chat, especially in large classrooms. Anecdotally, we noted that longer pauses may be necessary in a Zoom lecture compared to in-person lectures, as students may take extra time to both get up the courage to speak and to unmute themselves. We also saw the participation trends did change over time, for better or worse, so it is never too late in the semester for instructors to re-evaluate their pedagogical choices and encourage broad and diverse participation.

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

SUPPLEMENTAL FILE 1, PDF file, 1.1 MB.

ACKNOWLEDGMENTS

We acknowledge Brigham Young University for funding this study as part of a College Undergraduate Research Award given to author Sierra C. Nichols. We also acknowledge the assistance of Badí Quinteros Espinoza, who analyzed classroom data that we did not end up including in the paper.

REFERENCES

1. Martinez A, Christnacht C. 2021. Women are nearly half of U. S. workforce but only 27% of STEM workers, on United States Census Bureau. https://www.census.gov/library/stories/2021/01/women-making-gains-in-stem-occupations-but-still-underrepresented.html. Accessed September 30, 2021.
2. National Science Foundation. 2021. Women, minorities, and persons with disabilities in science and engineering: 2021. National Science Foundation, Alexandria, VA.
3. Wickware P. 1997. Along the leaky pipeline. Nature 390:202–203. https://doi.org/10.1038/36639.
4. Luckenbill-Edds L. 2002. The educational pipeline for women in biology: no longer leaking? Bioscience 52:513–521. https://doi.org/10.1641/0006-3568(2002)052[0513:TEPFWI]2.0.CO;2.
5. Goulden M, Mason MA, Frasch K. 2011. Keeping women in the science pipeline. Ann Am Acad Pol Soc Sci 638:141–162. https://doi.org/10.1177/0002716211416925.
6. West JD, Jacquet J, King MM, Correll SJ, Bergstrom CT. 2013. The role of gender in scholarly authorship. PLoS One 8: e66212. https://doi.org/10.1371/journal.pone.0066212.
7. Benish S. 2018. Meeting STEM workforce demands by diversifying STEM. J Science Policy & Governance 13:1–6.
8. Jehn KA, Northcraft GB, Neale MA. 1999. Why differences make a difference: a field study of diversity, conflict and performance in workgroups. Adm Sci Q 44:741–763. https://doi.org/10.2307/2667054.

9. Page SE. 2008. The difference: how the power of diversity creates better groups, firms, schools, and societies-new edition. Princeton University Press, Princeton, NJ.
10. Eddy SL, Brownell SE. 2016. Beneath the numbers: a review of gender disparities in undergraduate education across science, technology, engineering, and math disciplines. Phys Rev Phys Educ Res 12: e020106. https://doi.org/10.1103/PhysRevPhysEducRes.12.020106.
11. Guzzetti BJ, Williams WO. 1996. Gender, text, and discussion: examining intellectual safety in the science classroom. J Res Sci Teach 33:5–20. https://doi.org/10.1002/(SICI)1098-2736(199601)33:1%3C5::AID-TEA1%3E3.0.CO;2-Z.
12. Holmes J. 1992. Women's talk in public contexts. Discourse & Society 3:131–150. https://doi.org/10.1177/0957926592003002001.
13. Xu C, Lastrapes RE. 2021. Impact of STEM sense of belonging on career interest: the role of STEM attitudes. J Career Development. https://doi.org/10.1177/08948453211033025.
14. Athens W. 2018. Perceptions of the persistent: engagement and learning community in underrepresented populations. OLJ 22:27–58. https://doi.org/10.24059/olj.v22i2.1368.
15. Mendez G, Buskirk TD, Lohr S, Haag S. 2008. Factors associated with persistence in science and engineering majors: an exploratory study using classification trees and random forests. J Engineering Education 97:57–70. https://doi.org/10.1002/j.2168-9830.2008.tb00954.x.
16. Shedlosky-Shoemaker R, Fautch JM. 2015. Who leaves, who stays? Psychological predictors of undergraduate chemistry students' persistence. J Chem Educ 92:408–414. https://doi.org/10.1021/ed500571j.
17. Bailey E, Greenall R, Baek D, Morris C, Nelson N, Quirante T, Rice N, Rose S, Williams K. 2020. Female in-class participation and performance increase with more female peers and/or a female instructor in life sciences courses. CBE Life Sci Educ 19: ar30. https://doi.org/10.1187/cbe.19-12-0266.
18. Eddy SL, Brownell SE, Wenderoth MP. 2014. Gender gaps in achievement and participation in multiple introductory biology classrooms. CBE Life Sci Educ 13:478–492. https://doi.org/10.1187/cbe.13-10-0204.
19. Ballen CJ, Danielsen M, Jørgensen C, Grytnes J-A, Cotner S. 2017. Norway's gender gap: classroom participation in undergraduate introductory science. Njsteme 1:262–270. https://doi.org/10.5324/njsteme.v1i1.2325.
20. Ballen CJ, Aguillon SM, Awwad A, Bjune AE, Challou D, Drake AG, Driessen M, Ellozy A, Ferry VE, Goldberg EE, Harcombe W, Jensen S, Jørgensen C, Koth Z, McGaugh S, Mitry C, Mosher B, Mostafa H, Petipas RH, Soneral PAG, Watters S, Wassenberg D, Weiss SL, Yonas A, Zamudio KR, Cotner S. 2019. Smaller classes promote equitable student participation in STEM. Bioscience 69:669–680. https://doi.org/10.1093/biosci/biz069.
21. Nadile EM, Alfonso E, Barreiros BM, Bevan-Thomas WD, Brownell SE, Chin MR, Ferreira I, Ford SA, Gin LE, Gomez-Rosado JO, Gooding G, Heiden A, Hutt AE, King ML, Perez SG, Rivera Camacho YI, Salcedo F, Sellas CF, Sinda KA, Stahlhut KN, Stephens MD, Wiesenthal NJ, Williams KD, Zheng Y, Cooper KM. 2021. Call on me! Undergraduates' perceptions of voluntarily asking and answering questions in front of large-enrollment science classes. PLoS One 16:e0243731. https://doi.org/10.1371/journal.pone.0243731.
22. Wright CD, Eddy SL, Wenderoth MP, Abshire E, Blankenbiller M, Brownell SE. 2016. Cognitive difficulty and format of exams predicts gender and socioeconomic gaps in exam performance of students in introductory biology courses. CBE Life Sciences Education 15(2):ar23. https://doi.org/10.1187/cbe.15-12-0246.
23. Lauer S, Momsen J, Offerdahl E, Kryjevskaia M, Christensen W, Montplaisir L. 2013. Stereotyped: investigating gender in introductory science courses. CBE Life Sci Educ 12:30–38. https://doi.org/10.1187/cbe.12-08-0133.
24. Willoughby SD, Metz A. 2009. Exploring gender differences with different gain calculations in astronomy and biology. American J Physics 77:651–657. https://doi.org/10.1119/1.3133087.
25. Odom S, Boso H, Bowling S, Brownell S, Cotner S, Creech C, Drake AG, Eddy S, Fagbodun S, Hebert S, James AC, Just J, St Juliana JR, Shuster M, Thompson SK, Whittington R, Wills BD, Wilson AE, Zamudio KR, Zhong M, Ballen CJ. 2021. Meta-analysis of gender performance gaps in undergraduate natural science courses. CBE Life Sci Educ 20:ar40. https://doi.org/10.1187/cbe.20-11-0260.
26. Hodges C, Moore S, Lockee B, Trust T, Bond A. 2020. The difference between emergency remote teaching and online learning. Educause Rev 27:1–12.
27. Bashir A, Bashir S, Rana K, Lambert P, Vernallis A. 2021. Post-Covid-19 adaptations; the shifts towards online learning, hybrid course delivery and the implications for biosciences courses in the higher education setting. Front Edu 6.
28. Lee K, Fanguy M, Lu XS, Bligh B. 2021. Student learning during COVID-19: it was not as bad as we feared. Distance Education 42:164–172. https://doi.org/10.1080/01587919.2020.1869529.
29. Yusuf BN. 2020. Are we prepared enough? A case study of challenges in online learning in a private higher learning institution during the Covid-19 outbreaks. ASSRJ 7:205–212. https://doi.org/10.14738/assrj.75.8211.
30. Ferri F, Grifoni P, Guzzo T. 2020. Online learning and emergency remote teaching: opportunities and challenges in emergency situations. Societies 10:86. https://doi.org/10.3390/soc10040086.
31. Shin M, Hickey K. 2021. Needs a little TLC: examining college students' emergency remote teaching and learning experiences during COVID-19. J Further and Higher Education 45:973–986. https://doi.org/10.1080/0309877X.2020.1847261.
32. Gin LE, Guerrero FA, Brownell SE, Cooper KM. 2021. COVID-19 and undergraduates with disabilities: challenges resulting from the rapid transition to online course delivery for students with disabilities in undergraduate STEM at large-enrollment institutions. CBE Life Sci Educ 20:ar36. https://doi.org/10.1187/cbe.21-02-0028.
33. Perry BL, Aronson B, Pescosolido BA. 2021. Pandemic precarity: COVID-19 is exposing and exacerbating inequalities in the American heartland. Proc Natl Acad Sci U S A 118: e2020685118. https://doi.org/10.1073/pnas.2020685118.
34. Tsolou O, Babalis T, Tsoli K. 2021. The impact of COVID-19 pandemic on education: social exclusion and dropping out of school. CE 12:529–544. https://doi.org/10.4236/ce.2021.123036.
35. Castelli FR, Sarvary MA. 2021. Why students do not turn on their video cameras during online classes and an equitable and inclusive plan to encourage them to do so. Ecol Evol 11:3565–3576. https://doi.org/10.1002/ece3.7123.

36. Fauville G, Luo M, Muller Queiroz AC, Bailenson JN, Hancock J. 2021. Nonverbal mechanisms predict zoom fatigue and explain why women experience higher levels than men. SSRN https://doi.org/10.2139/ssrn.3820035.
37. Reinholz DL, Stone-Johnstone A, White I, Sianez LM, Jr, Shah N. 2020. A pandemic crash course: learning to teach equitably in synchronous online classes. CBE Life Sci Educ 19:ar60. https://doi.org/10.1187/cbe.20-06-0126.
38. Ashong CY, Commander NE. 2012. Ethnicity, gender, and perceptions of online learning in higher education. MERLOT J Online Learning and Teaching 8.
39. Johnson RD. 2011. Gender differences in e-learning: communication, social presence, and learning outcomes. J Organ End User Comput 23:79–94. https://doi.org/10.4018/joeuc.2011010105.
40. McSporran M, Young S. 2011. Does gender matter in online learning? Res in Learning Technology 9. https://doi.org/10.3402/rlt.v9i2.12024.
41. Xu D, Jaggars SS. 2014. Performance gaps between online and face-to-face courses: differences across types of students and academic subject areas. The J Higher Education 85:633–659. https://doi.org/10.1353/jhe.2014.0028.
42. Sullivan P. 2002. "It's easier to be yourself when you are invisible": female college students discuss their online classroom experiences. Innovative Higher Education 27:129–144. https://doi.org/10.1023/A:1021109410893.
43. Caspi A, Chajut E, Saporta K. 2008. Participation in class and in online discussions: gender differences. Computers & Education 50:718–724. https://doi.org/10.1016/j.compedu.2006.08.003.
44. Yaghmour S. 2012. Gender communication differences in synchronized distance learning lectures for users aged 18–60: a systematic literature review. ARECLS 9:85–90.
45. Mazzolini M, Maddison S. 2002. Does gender influence discussion forum participation in online education? p 421–430. Australasian Society for Computers in Learning in Tertiary Education Conference, Auckland, New Zealand.
46. Denaro K, Sato B, Harlow A, Aebersold A, Verma M. 2021. Comparison of cluster analysis methodologies for characterization of Classroom Observation Protocol for Undergraduate STEM (COPUS) data. Cbe-Life Sciences Education 20. https://doi.org/10.1187/cbe.20-04-0077.
47. Smith MK, Jones FH, Gilbert SL, Wieman CE. 2013. The Classroom Observation Protocol for Undergraduate STEM (COPUS): a new instrument to characterize university STEM classroom practices. CBE Life Sci Educ 12:618–627. https://doi.org/10.1187/cbe.13-08-0154.
48. Akiha K, Brigham E, Couch BA, Lewin J, Stains M, Stetzer MR, Vinson EL, Smith MK. 2018. What types of instructional shifts do students experience? Investigating active learning in science, technology, engineering, and math classes across key transition points from middle school to the university level. Front Educ 2:68. https://doi.org/10.3389/feduc.2017.00068.
49. Stains M, Harshman J, Barker MK, Chasteen SV, Cole R, DeChenne-Peters SE, Eagan MK, Esson JM, Knight JK, Laski FA, Levis-Fitzgerald M, Lee CJ, Lo SM, McDonnell LM, McKay TA, Michelotti N, Musgrove A, Palmer MS, Plank KM, Rodela TM, Sanders ER, Schimpf NG, Schulte PM, Smith MK, Stetzer M, Van Valkenburgh B, Vinson E, Weir LK, Wendel PJ, Wheeler LB, Young AM. 2018. Anatomy of STEM teaching in North American universities. Science 359:1468–1470. https://doi.org/10.1126/science.aap8892.
50. Grunspan DZ, Eddy SL, Brownell SE, Wiggins BL, Crowe AJ, Goodreau SM. 2016. Males under-estimate academic performance of their female peers in undergraduate biology classrooms. PLoS One 11:e0148405. https://doi.org/10.1371/journal.pone.0148405.
51. Moss-Racusin CA, Dovidio JF, Brescoll VL, Graham MJ, Handelsman J. 2012. Science faculty's subtle gender biases favor male students. Proc Natl Acad Sci U S A 109:16474–16479. https://doi.org/10.1073/pnas.1211286109.