Title
Effects of Remote Teaching in a Crisis on Equity Gaps and the Constructivist Learning Environment in an Introductory Biology Course Series

Journal
Journal of Microbiology and Biology Education, 22(1)

ISSN
1935-7877

Authors
Zuckerman, Austin L
Hardesty, Rebecca A
Denaro, Kameryn
et al.

Publication Date
2021-04-30

DOI
10.1128/jmbe.v22i1.2293

Peer reviewed

DOI: https://doi.org/10.1128/jmbe.v22i1.2293
JMBE Journal of Microbiology & Biology Education

Effects of Remote Teaching in a Crisis on Equity Gaps and the Constructivist Learning an Environment in an Introductory Biology Course Series†

Austin L. Zuckerman[1], Rebecca A. Hardesty[2], Kameryn Denaro[3], Stanley M. Lo[1], Melinda T. Owens[1]*
[1] Biological Sciences, University of California San Diego, La Jolla, CA 92093;
[2] National Cancer Institute, National Institutes of Health, Rockville, MD 20850;
[3] Division of Teaching Excellence and Innovation, University of California Irvine, Irvine, CA 92697

Because of the COVID-19 pandemic in March 2020, higher education institutions had to pivot rapidly to online remote learning. Many educators were concerned that the disparate impact of this crisis would exacerbate inequities in learning outcomes and student learning experiences, especially for students from minoritized backgrounds. We examined course grades and student perceptions of their learning experiences in fall (face-to-face) and spring (fully remote) quarters in an introductory biology course series at a public research university. Contrary to our hypothesis, we found that student course grades increased overall during remote learning, and equity gaps in course grades were mitigated for minoritized students. We hypothesize that instructors may have changed their grading practices to compensate for challenges in remote learning in crisis. However, spring students reported significant decreases in the amount of peer negotiation and social support, critical components of active learning. These findings suggest that remote teaching in crisis may have negatively affected student learning environments in ways that may not have been captured by grading practices.

INTRODUCTION

COVID-19 rapidly spread across the world and exerted various pressures on medical providers, national governments, and the general public (1, 2). Another sector that faced an unprecedented burden of adapting to the global health crisis was higher education. After the World Health Organization declared COVID-19 a pandemic in March 2020 (3), most colleges and universities transitioned to remote learning to slow the spread of the virus. Faculty and administrators had to quickly transition to remote learning with little preparation to adapt their curricula to an online platform, and students participating in remote learning were no longer able to access on-campus resources as they had in the past.

The quality of remote learning environments varies greatly across students, and issues related to internet access, familial obligations, and other personal responsibilities can be particularly disruptive to the cultivation of rigorous yet equitable learning experiences for all students (4). A survey implemented across multiple U.S. institutions reported that students faced increased struggles finding quiet spaces to work, completing collaborative and technical course assignments, and maintaining a sense of belonging within their institutions (5). Therefore, an increasing concern was the potential of the crisis to exacerbate existing equity gaps. Equity gaps are sustained disparities in educational attainment and learning outcomes between different groups of students (6–9), particularly minoritized students, including women, first-generation college students, and PEERs (persons excluded from science because of ethnicity or race) (10–12). Here, we use the term "equity gap" instead of "achievement gap" to highlight how these gaps arise from inequitable education systems, not inherent differences in student ability (13). Although there are anecdotal claims highlighting the increased academic hardships that students from minoritized backgrounds have faced during remote learning (14), there has yet to be any concrete research of how remote learning during a crisis has affected equity gaps in higher education. With many colleges and universities currently committing to full or partial remote learning, it is important to identify whether equity gaps continue to grow. With this knowledge, faculty and administrators can take action to improve remote teaching practices and prepare for future crises that may mandate a rapid transition to remote learning.

*Corresponding author. Mailing address: Biological Sciences, University of California San Diego, La Jolla, CA 92093. Phone: 858-534-7096. E-mail: mtowens@ucsd.edu.
Received: 30 September 2020, Accepted: 12 January 2021, Published: 31 March 2021
†Supplemental materials available at http://asmscience.org/jmbe

It is also essential to understand how remote learning has shifted student perceptions of their learning experiences. Exhaustive meta-analyses have demonstrated that pedagogical practices collectively referred to as "active learning" increase undergraduate academic performance, retention rates, and knowledge gains in STEM for all students (9, 15). Active-learning classrooms typically center around constructivist learning models. Constructivist learning practices support a shift from teacher-centered passive transmission of knowledge to a student-centered learning environment, where students engage in interactive peer discussions, connect with the course material at a personal level, learn about the uncertainty of science, and can share control and voice criticisms of their learning environment (16–18). As a result, these practices foster a community that provides instructor and peer support in learning (19). These practices are particularly important for retaining minoritized students (women, first-generation students, and PEERs), who have lower rates of persistence in STEM (20, 21). Understanding students' perceptions of the extent to which their experiences include these practices and how those perceptions have changed may help us predict how remote learning during a crisis may affect student learning gains in undergraduate STEM coursework.

We use courses from an introductory biology course series at a large research university on the quarter system as our sample. We compare the fully in-person instruction (fall) to fully remote instruction under crisis (spring). In our context, "remote learning" refers to courses taught entirely online with no face-to-face component. This study addresses the following research questions:

1. To what extent were equity gaps in course grades affected by remote learning under crisis?
2. To what extent did remote learning change student's perceptions of the learning experiences of their courses?

Because of inequities in the quality of remote learning environments, we expected that equity gaps in course grades would be maintained or exacerbated for students from minoritized backgrounds. We also hypothesized that because of the hasty switch away from face-to-face instruction and increased stress, students would experience less peer interaction and support, be less able to voice concerns or share control of their learning, not feel as connected to the material or science, and overall feel less supported in their learning. Therefore, we thought students would, on average, be less likely to report that they experienced constructivist learning environments associated with active learning.

TABLE 1.
Summary of demographics for each data set.

Demographics	Full Survey Sample (n = 2,460)	Propensity Matched Sample (n = 2,046)	Fall/Spring Paired Grades Sample (n = 184)	Fall/Spring Survey Sample[a] (n = 392)
Gender[b]				
Women	70%	70%	77%	79%
Men	30%	30%	23%	21%
Not Given	<1%			
PEER Status[c]				
PEER	34%	35%	40%	40%
Non-PEER	64%	65%	60%	60%
Not Given	3%			
College generation status				
First generation	34%	35%	39%	36%
Continuing generation	66%	65%	61%	64%
Major Class[d]				
Biological sciences	46%	48%	62%	60%
Biology-dependent	31%	30%	30%	31%
Non-biology	23%	22%	8%	8%

[a] Fall/spring survey sample consists of 196 unique students that completed a survey for a course in both the fall and spring quarters.

[b] Transgender individuals were grouped with their stated gender.

[c] Non-white or non-Asian individuals were classified as PEERs, and others were classified as non-PEERs.

[d] Students were also categorized by major into one of three classes: students with a major in the biology division (Biological sciences), students with a major that requires introductory biology coursework (Biology-dependent), and students with a major that does not require introductory biology coursework (Non-biology).

METHODS

Study context and participants

This study took place at a public, doctoral-granting university in the western United States with "very high research activity," as described by the Carnegie Classification of Institutions of Higher Education (22). All 29 courses in a four-course introductory biology series were invited to participate in a larger study examining university teaching and learning. The courses in this series are required for biology majors and many pre-health professions but may also be applied towards general education requirements for non-biology majors. Most of the courses do not need to be taken in a particular order, and some may be taken simultaneously. All courses had more than 100 students, with nearly all having 200 to 400 students.

At the end of the fall, winter, and spring quarters of the 2019–2020 academic year, students were asked to complete an online survey with items adapted from the Constructivist Learning Environment Survey (CLES) (18) and the Classroom and School Community Inventory (CSCI) (19) (Appendix 1). Survey completion was incentivized at the discretion of the instructor, with most instructors awarding a small amount of extra credit for completion. We received human subjects approval through the UC San Diego Project #191318XX.

In aggregate, 4,149 survey responses from 27 courses were collected across the three quarters. Demographics and grade data came from institutional records. Survey responses with missing demographic data (n = 98) were dropped.

Identifying appropriate comparison groups

All analyses were performed using the open-source programming environment R (R Foundation for Statistical Computing, Vienna, Austria [https://www.R-project.org/]). To analyze the effect that remote learning under a crisis had on student grade equity gaps and student perceptions of the learning environments, we first identified appropriate comparison groups. A natural experiment was identified, with fall 2019 serving as the control quarter compared with the spring 2020 crisis quarter. We did not use data from the winter quarter because the crisis was officially declared at the end of the winter quarter, and there were many significant transitions in course policies during the final 2 weeks of the term. In contrast, data from spring quarter were used because all coursework and examinations were implemented remotely and all student course experiences occurred amidst the crisis. We received 2,460 responses from students in fall and spring after removing students who did not agree to fully participate and release their demographic information and educational records (Table 1).

A fully randomized control was not available for this study, so we identified appropriate comparison groups between fall and spring in two ways. First, we identified a "paired sample" consisting of individual students who completed the survey in both the fall and spring quarters and took an introductory biology course for a letter grade each quarter. Because this sample was made up of the same individuals, differences in their grades or survey scores in spring compared to fall are more likely to result from learning during the crisis. We used this sample to analyze equity gaps and student perceptions of their learning environment. However, this sample was very small, consisting of only 196 students (392 responses), of whom only 184 students (368 responses) took the course for a letter grade in both terms. Therefore, we created a propensity score–matched sample for fall and spring. This method mimics a randomized controlled trial to correct for sample selection bias due to possible differences in subject characteristics between the control (fall) and treatment (spring) groups (23). We used the propensity score–matched comparison groups to analyze differences in equity gaps in fall and spring because it maintains a large sample size, which is necessary for detecting significant but small effects. This method resulted in 1,023 individuals that completed a course in fall for a letter grade and 1,023 individuals that completed a course in spring for a letter grade.

To generate matched control and treatment groups, nearest-neighbor matching was used to generate the propensity scores using the MatchIt package (24). The following covariates were used to create the propensity-matched dataset: gender (classified into men or women), first-generation college status (first-generation or continuing generation college student), ethnicity, course, major, cumulative units completed at the end of the fall quarter, and fall quarter cumulative grade point average excluding introductory biology courses during the fall quarter (adjGPA) (25, 26).

Analyzing equity gaps

To examine if remote learning during the crisis had altered grades across different demographic groups, we used linear mixed effects (LME) models. Course grades were converted to a 4.0 scale, and credit/no credit, incompletes, withdrawals, and blank grades were omitted from the data set. Aligning with the institution's grading policies, both A and A+ grades were assigned a 4.0. One model was created for the propensity-matched sample and one for the paired sample. The propensity-matched sample was used to compare groups of students who took courses in fall and spring that were matched for a variety of characteristics, as detailed in the section "Identifying appropriate comparison groups," above. The paired sample was used to compare the performance of the exact same students who took courses in both fall and spring to eliminate the effects of differences between individuals who took courses in various quarters. Details of the model selection are given in the "Model selection" section.

Measuring and analyzing student learning experiences

Student perceptions of their classroom learning environments were examined using the five affective subscales from the CLES and the two subscales from the CSCI. The survey (reprinted in Appendix 1) was accessible during the last 2 weeks of the quarter. There were six items for each subscale from the CLES, and each item was evaluated on a five-point frequency scale (Table 2). There were five items each for the additional two subscales from the CSCI. For consistency, items in these two subscales were rated on a six-point Likert-type scale so that the maximum sum of all items within each of the seven subscales was 30. A 6-point scale was used to avoid having an ambiguous midpoint option, which could be interpreted as "neutral" or "undecided," two similar but distinct constructs (27–29). To assess the reliability of our survey, we calculated Cronbach's alpha to measure the intercorrelation of responses to the items hypothesized to be related to each subscale (30). The Cronbach's alpha for most survey subscales was above 0.75, indicating high internal consistency within the survey items for each subscale (Table 2).

To obtain a structured comparison of how responses to these scales shifted from fall to spring, we generated seven separate LME models for each of the seven subscales with the composite survey score for each subscale as the dependent variable. We used a paired sample consisting of survey responses from 196 students who completed the survey in both fall and spring to increase the chance that any changes in survey scores due to quarter were due to learning under crisis and not differences in individuals who took the survey. For students who happened to take multiple introductory biology courses in a single quarter and thus completed the survey for multiple courses, one of the survey responses was randomly selected for analysis. Details of the model creation and selection are given in the section "Model creation and selection," below.

Mathematics of LME regression models

LME regression models are similar to ordinary least squares (OLS) regression models because a researcher can model the response (y) as a linear combination of p predictors ($\chi = \chi_1, \cdots \chi_p$). However, the linear mixed effects regression models also account for the correlation between students. In our case, for example, we know that students taking the same course section are probably more similar to each other than they are to students taking a different course. These correlations between students would violate the assumptions of the OLS regression model, whereas the LME model does not assume independent observations (31, 32). The linear mixed model is given by:

$$y_{ij} = x_{ij}^T\beta + u_{ij}^T\gamma_i + \varepsilon_{ij}$$

where y_{ij} is the response of the jth student of class i ($i = 1,\ldots, 4$, $j = 1,\ldots, n_i$), n_i is the size of the class i, x_{ij} is the covariate vector of the jth student of class i for the fixed effects (quarter, adjGPA from fall quarter, whether or not a student is a biology major, gender, first-generation college status, and PEER status), β is the fixed effects parameter, u_{ij} is the covariate vector of the jth student of class i for the random effects, γ_i is the random effect parameter, ε_{ij} is the random error associated with the jth student of class i, and ε_i is the error

TABLE 2.
Definitions and sample items for each survey subscale.

Source[a]	Survey Subscale	Definition	Sample Item	Cronbach's Alpha
CLES	Personal Relevance	Relevance of learning to students' lives	I learn about the world outside of school.	0.786
CLES	Uncertainty of Science	Provisional status of scientific knowledge	I learn that science has changed over time.	0.674
CLES	Critical Voice	Legitimacy of expressing a critical opinion	It is okay for me to question the way that I'm being taught.	0.802
CLES	Shared Control	Participating in planning, conducting and assessing of learning	I help the instructor to plan what I am going to learn.	0.904
CLES	Peer Negotiation	Discussing ideas with other students	I get the chance to talk to other students.	0.942
CSCI	Social Support	Feelings of community regarding cohesion, trust, interdependence, and sense of belonging	I feel connected to others in the course.	0.906
CSCI	Learning Support	Sharing group norms and values; the extent to which educational goals and expectations are met	I feel that my educational needs are being met in the course.	0.795

[a] Survey questions were derived from either the CLES (18) or the classroom form of the CSCI (19).

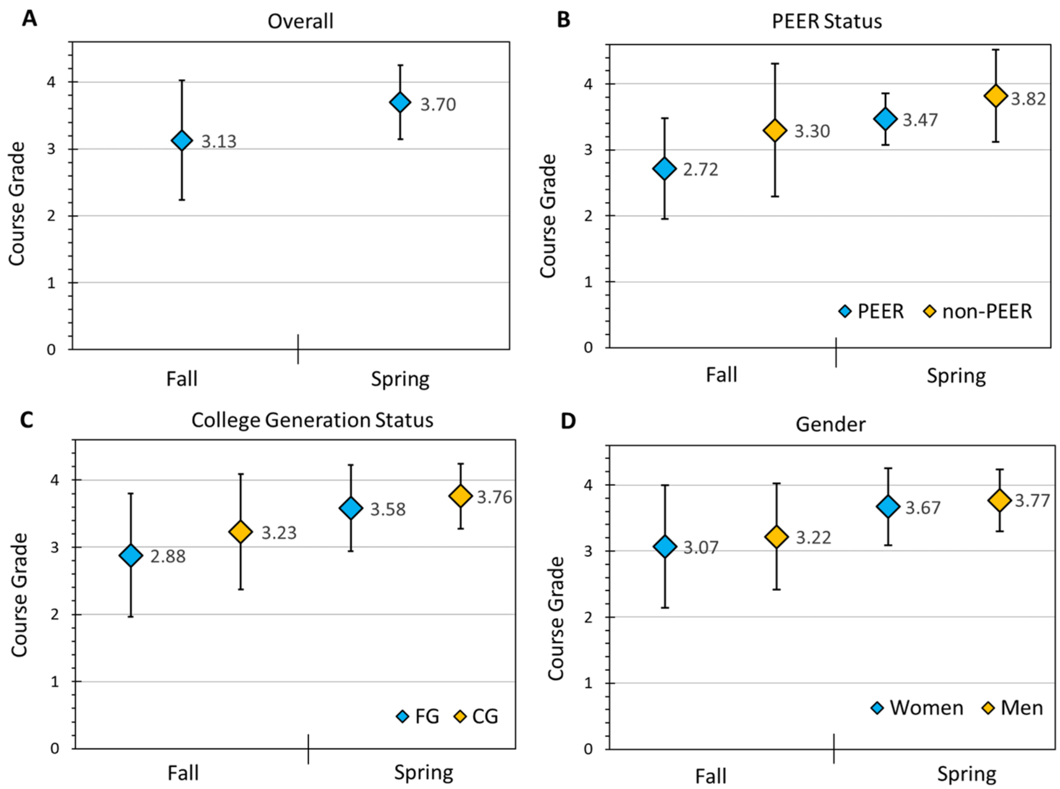

FIGURE 1. Comparison of average course grades in fall and spring quarters for all propensity-matched students (A), PEERs vs. non-PEERS (B), first-generation (FG) vs. continuing-generation (CG) college students (C), and women vs. men (D). Error bars indicate standard deviations.

vector of class *i*. The model assumptions are: (i) the random effects parameter follows a Gaussian (normal) distribution with mean zero and covariance matrix D; (ii) the random error for class *i* follows a Gaussian distribution with mean zero and covariance matrix Σ_i, and (iii) each of the random effect parameters and random errors are independent. For a complete discussion of LME regression models, see Laird and Ware (33). Equity gaps in biology are modeled as a linear combination of the student-level covariates and the random error representing the influence of class *i* on the student that is not captured by the observed covariates. The random cluster errors are added to the regression model to account for the correlation of the students within each biology class.

Model creation and selection

LME models were created using the lme4 package (34). For each LME model created, we initially added quarter, gender, first-generation college-going status, PEER status, major class, and adjusted GPA from fall quarter as possible fixed effects. "PEERs" were defined as students who were non-white and non-Asian. "Major class" refers to whether a student's major is in the biology division (Biological sciences), is not biology but requires introductory biology coursework (Biology-dependent), or is a non-biology major that does not require introductory biology coursework (Non-biology). For modeling equity gaps, we sought to determine if the course grades of different demographic groups were disproportionately affected by the transition to remote learning. Therefore, we also added Quarter and PEER status, Quarter and gender, and Quarter and first-generation college status as possible interaction effects for those models. For all models, possible random effects (random intercept only) included course, course section, and instructor. In models that used the paired sample, individual student was also specified as a possible random effect.

The lmer and lmerTest packages were used to select the LME models (34, 35). All models were fit using restricted maximum likelihood (REML). The *p* values were calculated using *t* tests with Satterthwaite's method. Following Theobald (2018), to select random effects for the final model, we generated models with all possible fixed effects and all combinations of random effects and compared them using the Akaike information criteria (AICc) with a penalty for small sample sizes (32). AICc provides a relative goodness-of-fit test for models, with the lowest AICc indicating the best-fitting model (32). Then, to select fixed effects, we used the dredge function in the MuMIn package (version 1.43.17; K Bartoń [https://cran.r-project.org/web/packages/MuMIn/index.html]) to test all combinations of possible fixed effects. The models with the lowest AICc values were compared using analysis of variance. For models with small differences in AICc (<2) that were not significantly different, the more parsimonious model was used.

RESULTS

Course grades and equity gaps in remote learning under crisis

We quantified the impact of remote learning on course grades and equity gaps by comparing course grades earned in fall 2019 and spring 2020. Among all students, the spring grades were much higher, with an average grade of 3.13 in fall and 3.70 in spring (Fig. 1A) and respective standard deviations of (±0.89) and (±0.55). To further explore how course grades were affected by quarter and factors like membership in minoritized groups (PEERs, women, first-generation students), we created an LME model to account for the variation in course grades that was associated with these factors. First, to select comparable groups of fall and spring quarter students, we used propensity scoring to match students that completed coursework in fall and spring quarters (n = 1,023 students per quarter). Then, we selected the "best-fitting" LME model, the one with the combination of variables that yields the lowest AICc (corrected for small sample sizes). AICc is a metric that balances significance and parsimony to choose which factors contribute significantly to variation in course grades (32). A summary of the best-fitting LME model and effect size estimates for each of its fixed effects are summarized in Table 3. We set the significance level (to 0.05; variables with p-values smaller than the predetermined significance level are considered to be significant. Effect size estimates are listed as β. These represent the change in course grades associated with that variable. For example, the β for Quarter (SP20) was 0.469, meaning that if we compared students who completed remote coursework in spring quarter to students who took conventional courses in fall, the spring students' course grades would have been, on average, 0.469 GPA points higher, if all the other variables in the model (course section, fall GPA, major class, PEER status, etc.) were held constant (Table 3). The increase in grades associated with spring quarter was significant ($p < 0.01$).

We found that students from minoritized groups (PEERs, women, first-generation students) had lower grades on average, as the β_{PEER}, β_{Women}, and β_{FG} were all negative, consistent with much previous literature (7, 36, 37) (Table 3). However, contrary to our expectations, the model shows a smaller equity gap for PEERs compared to non-PEERs and first-generation students compared to continuing-generation students in spring than in fall (Fig 1B and C). In addition to the main effects for quarter and minoritized groups (β_{Spring}, β_{PEER}, β_{Women}, and β_{FG}), we considered that there could have been a differential effect of remote learning for minoritized students compared to non-minoritized students, and we looked for that by including interaction terms between quarter and group ($\beta_{Spring*PEER}$, $\beta_{Spring*Women}$, and $\beta_{Spring*FG}$) as potential fixed effects. To understand these interaction terms and how they relate to equity, let us start with the estimated equity gap for PEERs. In our final model, $\beta_{Spring*PEER}$ is 0.128. That means the estimated equity gap in spring for PEERs is only –0.125 GPA points ($\beta_{PEER} + \beta_{Spring*PEER}$ = -0.253 + 0.128), while in fall the estimated equity gap for PEERs is –0.253 (β_{PEER}), holding all other variables in the model constant (Table 3). Therefore, we can conclude that there were different effects in spring on the PEER and non-PEER groups, and the equity gap became smaller (Fig. 1B). Similarly, in our final model, the equity gap

TABLE 3.
Summary of final LME model for course grades using a propensity-matched student sample and a random effect for course section.[a]

Fixed Effects	Estimate (β)	SE	t value	p value
Intercept	1.678	0.119	14.100	<0.001
Quarter (SP20)	0.469	0.128	3.672	0.0018
Major Class (Bio-dependent)	-0.088	0.031	-2.809	0.0050
Major Class (Non-biology)	-0.040	0.034	-1.171	0.2416
Gender (Women)	-0.090	0.027	-3.232	0.0013
PEER Status (PEER)	-0.253	0.041	-6.188	<0.001
First Gen. Status (FG)	-0.093	0.040	-2.338	0.0195
Fall Quarter GPA	0.531	0.021	25.211	< 0.001
Quarter (SP20) * PEER Status (PEER)	0.128	0.056	2.298	0.0216
Quarter (SP20) * First Gen. Status (FG)	0.111	0.057	2.00	0.0456

[a] Initially, we included the following as possible fixed effects: quarter, major class, gender, first-generation status, PEER status, and the interaction between quarter and gender, first-generation status, and PEER status. The effects in the final model, chosen through the process described in "Model Selection," are included. n = 2,046 responses (1,023 students per quarter).

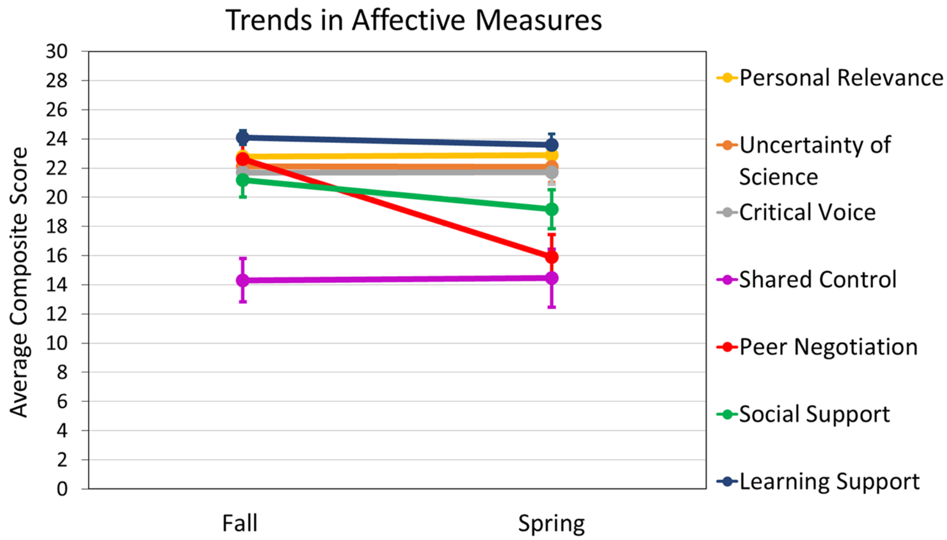

FIGURE 2. Affective measures from survey. For fall and spring courses, the average composite score (maximum of 30) for each of the seven affective subscales are plotted (n = 2,460 total responses). Error bars indicate standard deviations.

for first-generation compared to continuing-generation students was eliminated in spring (Fig. 1C). The estimated equity gap in fall for first-generation students is –0.093 (β_{FG}), and in spring is 0.018 ($\beta_{FG} + \beta_{Spring*FG} = -0.093 + 0.111$), holding all other variables in the model constant (Table 3). The final model chosen based on AICc only included the interactions terms between quarter and PEER status ($\beta_{Spring*PEER}$) and between quarter and FG status ($\beta_{Spring*FG}$). The interaction term for women and spring quarter ($\beta_{Spring*Women}$) was not found to be significant and thus was not included in the final model, suggesting that the equity gap for women did not change in spring quarter (Fig. 1D).

To exclude the possibility that increased grades in spring were because of changes in student populations between fall and spring, we generated a second LME model using data only from the students who completed both a fall and spring course (Table S1). This data set was much smaller, only 184 students. Even with a much smaller sample size, we again found that students also achieved significantly higher course grades in the spring quarter ($\beta = 0.493$, $p < 0.01$). We also found that PEER students had lower grades in this subsample ($\beta = -0.173$, $p < 0.05$), but we did not observe the smaller equity gaps for gender and college generation status, nor did we see interaction effects with quarter and minoritized groups.

Changes in the constructivist learning experience

To explore preliminary trends in the constructivist learning experience, we calculated the average composite survey score across each of the seven subscales for fall and spring for all students. For most subscales, the average composite score for each subscale was relatively constant between the two quarters, but there was a decrease in student perceptions of classroom social support and a substantial decrease (approximately 6 out of 30 points) in peer negotiation in the spring quarter (Fig. 2).

We then used LME models to compare constructivist learning experience outcomes for students (n = 196) that had completed the survey in both the fall and spring quarters (Table 4). Only the outcomes of peer negotiation, which measures how much students perceive that they discuss ideas with other students, and social support, which relates to how connected students feel with each other, included "quarter" as a significant effect in the final model. Students completing coursework in spring reported a drastic decrease in their sense of negotiation with peers in the learning process ($\beta = -6.482$, $p < 0.001$) and a smaller but still significant decrease in their sense of classroom social support ($\beta = -1.736$, $p < 0.01$). There was no significant effect of quarter for any other subscale.

DISCUSSION

This is one of the first studies that examines equity gaps and student learning experiences to understand how remote learning during the COVID-19 crisis impacted students. Because disparities in the impact of COVID-19 and the quality of remote learning environments may disadvantage students from minoritized backgrounds, we expected that the transition to remote learning would exacerbate equity gaps in grades for PEERs and first-generation students (4) and worsen constructivist learning experiences. However, contrary to our original hypothesis, we found that student grades were significantly higher in spring quarter (Table 3 and Table S1) and equity gaps were mitigated for PEERs and first-generation students (Table 3). We also found that perceptions of peer negotiation and social support were the only components of constructivist learning experiences disrupted in spring (Table 4).

Although it is possible that students simply performed better with online learning, the significant increase in course

TABLE 4.
Summary of linear-mixed effects models for each survey subscale.[a]

Survey Subscale	Fixed Effects	Estimate (β)	SE	t value	p value
Personal Relevance	Intercept	23.078	0.529	43.64	<0.001
Uncertainty of Science	Intercept	22.208	0.390	56.93	<0.001
Critical Voice	Intercept	21.839	0.331	66.02	<0.001
Shared Control	Intercept	15.110	0.624	24.23	<0.001
Peer Negotiation	Intercept	21.646	1.365	15.859	<0.001
	Quarter (SP20)	-6.482	0.594	-10.915	<0.001
	Major Class (Bio dep.)	-1.617	0.624	-2.593	0.0101
	Major Class (non-bio)	-1.362	1.043	-1.306	0.1927
	Fall GPA	0.784	0.386	2.031	0.0436
Social Support	Intercept	17.917	1.238	14.467	<0.001
	Quarter (SP20)	-1.736	0.527	-3.293	0.0015
	Fall GPA	1.250	0.364	3.434	<0.001
Learning Support	Intercept	20.843	1.005	20.751	<0.001
	First Gen. Status (FG)	1.196	0.465	2.573	0.0108
	Fall GPA	0.976	0.298	3.275	0.0013

[a] Survey data from students who completed survey in both fall and spring were analyzed (n = 392 responses from 196 students). For personal relevance, random effects were ID and course, while for all other subscales, random effects were ID and instructor. Initially, we included the following as possible fixed effects: quarter, major class, gender, first-generation status, PEER status, and fall GPA. The effects in the final model, chosen through the process described in "Model Selection," are included. Survey subscales with quarter as a significant fixed effect are bolded.

grades during the spring is likely attributable to changes in instructor grading practices, such as different grading policies in spring 2020 in reaction to the crisis. Education leaders in the biology division informally advised course instructors to be more lenient with deadlines and missed work and to have more frequent but smaller assessments. After the widespread Black Lives Matter protests over police brutality in late May 2020, campus leaders recognized the potential for protests to disparately decrease students' abilities to complete course requirements and strongly recommended that instructors implement final exams that could only improve student grades. These policies would tend to increase grades and decrease equity gaps. In addition, the decrease in equity gaps may come in part from a ceiling effect on grades, as the average grade in all 10 courses in our spring 2020 dataset was above a 3.4 out of 4.0.

These results raise questions about how instructors should interpret these spring 2020 grades. Studies have found that grades do not necessarily provide feedback to students, motivate students to learn, or correlate with student learning (38). So, it is important to recognize that increases in course grades and decreases in equity gaps do not necessarily reflect increased cognitive learning gains. That caution is especially warranted if increases in course grades arise from changes in grading policies or grading curves or if decreases in equity gaps come from a ceiling effect (39). However, it is known that grades do correlate with systemic inequities in the opportunities offered to students in different groups, which becomes problematic when grades are used to exclude students from opportunities such as continuing in their chosen major (40, 41). The policies cited above that may have contributed to higher grades were proposed to mitigate the effects of inequitable remote learning environments. Simply accommodating students' individual circumstances through changing grading practices decreased equity gaps and may have allowed students to overcome barriers to course completion that might have otherwise been insurmountable. This suggests that perhaps typical grading practices are more likely to ignore these circumstances and create larger equity gaps.

We originally hypothesized that the transition to remote learning would lead to decreased affective outcomes on all survey subscales. However, only student perceptions of peer negotiation and classroom social support significantly decreased in spring, both of which relate to student-student interaction. The other five subscales are either directly related to course content or implementation of course policies, which could be similar in either an in-person or remote classroom environment (Table 2).

Student collaboration with peers and engagement with diverse perspectives are considered key elements of the constructivist learning model (17). Many studies have shown that pedagogical approaches that include students talking to each other increase learning gains and decrease equity gaps (9, 15). Student collaboration would logically also lead to increased sense of community, which has been shown to contribute to increased retention in STEM, especially for students from minoritized groups (42–44). Therefore, we have reason to believe that this quarter of remote learning under crisis may be less effective for students' academic learning despite higher grades (45, 46). These results highlight the importance of increased professional development for instructors and technical support in online curriculum development to bolster student-student interaction and community in remote courses to further close equity gaps and improve learning, especially as the crisis continues (47).

There are several limitations to this study. First, different course sections did not have common learning objectives or assessments. Therefore, we cannot determine whether increases in grades or decreased perceptions of constructivist learning were associated with changes in learning outcomes. Second, because of low sample sizes, we regrettably had to group together students with distinct racial/ethnically-based experiences and issues into the category of PEERs (48). For example, it is reasonable to believe that Black students may have been uniquely affected by issues revealed by the Black Lives Matter protests. Third, we did not ask students about their remote learning environments or how COVID-19 impacted them and their families. Future studies that explore other information-rich datasets (e.g., student interviews or in-depth surveys) would create a more comprehensive picture of how the crisis quarter affected the experiences of different individuals, including those aggregated together in the PEER group. Finally, our findings are not intended to be generalizable. Instructors at our institution may have responded to the challenges of remote teaching under crisis in ways different from instructors at other institutions. However, our findings make it clear that institutions cannot merely look at student grades or even equity gaps to assess student experiences or learning outcomes.

With the uncertainty of the duration and impact of the COVID-19 pandemic and the everlasting potential for future crises to disrupt in-person learning, we encourage faculty and administrators to consider teaching strategies that facilitate increased interaction between students during periods of remote learning (49, 50). These strategies are fundamental for sustaining a constructivist learning environment where students are autonomous and accountable in their learning experiences (51). Additionally, we encourage faculty to use multiple forms of assessment to measure student learning outcomes. These assessment strategies can include concept inventories or other formative assessments that provide objective estimations of students' understanding of course content. Our study suggests that course grades may be an incomplete or inaccurate representation of student learning. We found decreased equity gaps and higher grades in spring despite a decrease in student perceptions of peer negotiation and classroom social support, which have previously been associated with learning gains. Therefore, it is essential for faculty to be flexible and creative when adapting their curricula to ensure continuity in learning during periods of crisis.

SUPPLEMENTAL MATERIALS

Appendix 1: Survey administered to students
Appendix 2: Table S1

ACKNOWLEDGMENTS

We are thankful to all the course instructors and students who participated in this study. This material is based upon work supported by the National Science Foundation, DUE Grant 1821724, and UC San Diego General Campus Research Award RG095169. The data in this article was collected while Rebecca Hardesty was employed at the University of California San Diego. The opinions expressed in this article are the authors' own and do not reflect the views of the National Institutes of Health, the Department of Health and Human Services, or the U.S. Government. The authors declare no conflict of interest.

REFERENCES

1. Wang C, Horby PW, Hayden FG, Gao GF. 2020. A novel coronavirus outbreak of global health concern. Lancet 395:470–473.
2. Kang L, Li Y, Hu S, Chen M, Yang C, Yang BX, Wang Y, Hu J, Lai J, Ma X, Chen J, Guan L, Wang G, Ma H, Liu Z. 2020. The mental health of medical workers in Wuhan, China dealing with the 2019 novel coronavirus. Lancet Psychiatry 7:e14.
3. World Health Organization. 2020. Coronavirus disease 2019 situation report 51, 11 March 2020.
4. Robinson L, Schulz J, Khilnani A, Ono H, Cotten SR, Mcclain N, Levine L, Chen W, Huang G, Casilli AA, Tubaro P. 2020. Digital inequalities in time of pandemic: COVID-19 exposure risk profiles and new forms of vulnerability. First Monday 25. doi:10.5210/fm.v25i7.10845.
5. Blankstein M, Frederick JK, Wolff-Eisenberg C. 2020. Student experiences during the pandemic pivot. Ithaka S+ R doi:10.18665/sr.313461.
6. Bensimon EM. 2005. Closing the achievement gap in higher education: an organizational learning perspective. New Dir High Educ 2005:99–111.
7. Sirin SR. 2005. Socioeconomic status and academic achievement:

a meta-analytic review of research. Rev Educ Res 75:417–453.
8. Jury M, Smeding A, Stephens NM, Nelson JE, Aelenei C, Darnon C. 2017. The experience of low-SES students in higher education: psychological barriers to success and interventions to reduce social-class inequality. J Soc Issues 73:23–41.
9. Theobald EJ, Hill MJ, Tran E, Agrawal S, Nicole Arroyo E, Behling S, Chambwe N, Cintrón DL, Cooper JD, Dunster G, Grummer JA, Hennessey K, Hsiao J, Iranon N, Jones L, Jordt H, Keller M, Lacey ME, Littlefield CE, Lowe A, Newman S, Okolo V, Olroyd S, Peecook BR, Pickett SB, Slager DL, Caviedes-Solis IW, Stanchak KE, Sundaravardan V, Valdebenito C, Williams CR, Zinsli K, Freeman S. 2020. Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. Proc Natl Acad Sci USA 117:6476–6483.
10. Asai DJ. 2020. Race matters. Cell 181:754–757.
11. Flack CB, Walker L, Bickerstaff A, Margetts C. 2020. Socioeconomic disparities in Australian schooling during the COVID-19 pandemic. Pivot Professional Learning, Melbourne, Australia.
12. Malisch JL, Harris BN, Sherrer SM, Lewis KA, Shepherd SL, McCarthy PC, Spott JL, Karam EP, Moustaid-Moussa N, Calarco JMC, Ramalingam L, Talley AE, Cañas-Carrell JE, Ardon-Dryer K, Weiser DA, Bernal XE, Deitloff J. 2020. Opinion: in the wake of COVID-19, academia needs new solutions to ensure gender equity. Proc Natl Acad Sci USA 117:15378–15381.
13. Ladson-Billings G. 2006. 2006 presidential address from the achievement gap to the education debt: understanding achievement in U.S. schools. Educ Res 35:3–12.
14. Fain P. 2020. Higher education and work amid crisis. Inside High Educ https://www.insidehighered.com/news/.
15. Freeman S, Eddy SL, McDonough M, Smith MK, Okoroafor N, Jordt H, Wenderoth MP. 2014. Active learning increases student performance in science, engineering, and mathematics. Proc Natl Acad Sci USA 111:8410–8415.
16. Cirik, Dİ, Çolak, E, Kaya, D. 2015. Constructivist learning environments: the teachers' and students' perspectives. Int J New Trends Educ Their Implic 6:30–44.
17. Hwa SP. 2020. Assessing university students' perceptions toward constructivist learning environments using the CLES. Int J Adv Res Educ Soc 2:115–129.
18. Nix RK, Fraser BJ, Ledbetter CE. 2005. Evaluating an integrated science learning environment using the constructivist learning environment survey. Learn Environ Res 8:109–133.
19. Rovai AP, Wighting MJ, Lucking R. 2004. The classroom and school community inventory: development, refinement, and validation of a self-report measure for educational research. Internet Higher Educ 7:263–280.
20. National Science Foundation, National Center for Science and Engineering Statistics. 2019. Women, minorities, and persons with disabilities in science and engineering: 2019. Special Report NSF 19-304, Alexandria, VA.
21. Dika SL, D'Amico MM. 2016. Early experiences and integration in the persistence of first-generation college students in STEM and non-STEM majors. J Res Sci Teach 53:368–383.
22. McCormick AC, Zhao CM. 2005. Rethinking and reframing the Carnegie classification. Chang Mag High Learn 37:51–57.
23. Austin PC. 2011. An introduction to propensity score methods for reducing the effects of confounding in observational studies. Multivariate Behav Res 46:399–424.
24. Ho DE, King G, Stuart EA, Imai K. 2011. MatchIt: nonparametric preprocessing for parametric causal inference. J Stat Softw 42:1–28.
25. Rosenbaum PR, Rubin DB. 1983. The central role of the propensity score in observational studies for causal effects. Biometrika 70:41–55.
26. Randolph JJ, Falbe K, Manuel AK, Balloun JL. 2014. A step-by-step guide to propensity score matching in R. Pract Assess Res Eval 19:1–6.
27. Armstrong RL. 1987. The midpoint on a five-point Likert-type scale. Percept Mot Skills 64:359–362.
28. Guy RF, Norvell M. 1977. The neutral point on a Likert scale. J Psychol 95:199–204.
29. Komorita SS. 1963. Attitude content, intensity, and the neutral point on a Likert scale. J Soc Psychol 61:327–334.
30. Taber KS. 2018. The use of Cronbach's alpha when developing and reporting research instruments in science education. Res Sci Educ 48:1273–1296.
31. Theobald R, Freeman S. 2014. Is it the intervention or the students? Using linear regression to control for student characteristics in undergraduate STEM education research. CBE Life Sci Educ 13:41–48.
32. Theobald E. 2018. Students are rarely independent: when, why, and how to use random effects in discipline-based education research. CBE Life Sci Educ 17:1–12.
33. Laird NM, Ware JH. 1982. Random-effects models for longitudinal data. Biometrics 38:963–974.
34. Bates D, Mächler M, Bolker BM, Walker SC. 2015. Fitting linear mixed-effects models using lme4. J Stat Softw 67:1–48.
35. Kuznetsova A, Brockhoff PB, Christensen RHB. 2017. lmerTest package: tests in linear mixed effects models. J Stat Softw 82:1–26.
36. Stephens NM, Hamedani MYG, Destin M. 2014. Closing the social-class achievement gap: a difference-education intervention improves first-generation students' academic performance and all students' college transition. Psychol Sci 25:943–953.
37. Matz RL, Koester BP, Fiorini S, Grom G, Shepard L, Stangor CG, Weiner B, McKay TA. 2017. Patterns of gendered performance differences in large introductory courses at five research universities. AERA Open 3:1–12.
38. Schinske J, Tanner K. 2014. Teaching more by grading less (or differently). CBE Life Sci Educ 13:159–166.
39. Aviles CB. 2001. Grading with norm-referenced or criterion-referenced measurements: to curve or not to curve, that is the question. Soc Work Educ 20:603–608.
40. Harris RB, Mack MR, Bryant J, Theobald EJ, Freeman S. 2020. Reducing achievement gaps in undergraduate general chemistry could lift underrepresented students into a "hyperpersistent zone." Sci Adv 6:1–9.
41. Barr DA, Gonzalez ME, Wanat SF. 2008. The leaky pipeline: factors associated with early decline in interest in premedical studies among underrepresented minority undergraduate students. Acad Med 83:503–511.

42. Ranga JS. 2020. Online engagement of commuter students in a general chemistry course during COVID-19. J Chem Educ doi:10.1021/acs.jchemed.0c00633.
43. Strayhorn TL. 2018. College students' sense of belonging: a key to educational success for all students. Routledge, New York, NY.
44. Rainey K, Dancy M, Mickelson R, Stearns E, Moller S. 2018. Race and gender differences in how sense of belonging influences decisions to major in STEM. Int J STEM Educ 5. doi:10.1186/s40594-018-0115-6.
45. Johnson VE 2003. Grade inflation: a crisis in college education. Springer Verlag, New York, NY.
46. Miller G. 2014. Grade inflation, gatekeeping, and social work education: ethics and perils. J Soc Work Values Ethics 11:12–22.
47. Kebritchi M, Lipschuetz A. 2017. Issues and challenges for teaching successful online courses in higher education: a literature review. J Educ Technol Syst 46:4–29.
48. Bensimon EM. 2017. The misbegotten URM as a data point. Center for Urban Education, Rossier School of Education, University of Southern California, Los Angeles, CA.
49. Dixson MD. 2010. Creating effective student engagement in online courses: what do students find engaging? J Scholarsh Teach Learn 10:1–13.
50. Abrami PC, Bernard RM, Bures EM, Borokhovski E, Tamim RM. 2011. Interaction in distance education and online learning: using evidence and theory to improve practice. J Comput High Educ 23:82–103.
51. Rovai AP. 2004. A constructivist approach to online college learning. Internet High Educ 7:79–93.