

Closing the Achievement Gap in a Large Introductory Course by Balancing Reduced In-Person Contact with Increased Course Structure

Sat Gavassa,^{†*} Rocio Benabentos,[‡] Marcy Kravec,[†] Timothy Collins,[†] and Sarah Eddy^{†*}

[†]Department of Biological Sciences and [‡]STEM Transformation Institute, College of Arts, Sciences & Education, Florida International University, Miami, FL 33199

ABSTRACT

Hybrid and online courses are gaining attention as alternatives to traditional face-to-face classes. In addition to the pedagogical flexibility afforded by alternative formats, these courses also appeal to campuses aiming to maximize classroom space. The literature, however, reports conflicting results regarding the effect of hybrid and online courses on student learning. We designed, taught, and assessed a fully online course (100% online) and a hybrid-and-flipped course (50% online 50% face-to-face) and compared those formats with a lecture-based face-to-face course. The three formats also varied in the degree of structure; the hybrid course was the most structured and the face-to-face course was the least structured. All three courses were taught by the same instructor in a large Hispanic-serving research university. We found that exam scores for all students were lowest in the face-to-face course. Hispanic and Black students had higher scores in the hybrid format compared with online and face-to-face, while white students had the highest performance in the online format. We conclude that a hybrid course format with high structure can improve exam performance for traditionally underrepresented students, closing the achievement gap even while in-person contact hours are reduced.

INTRODUCTION

To boost scientific discovery and grow the scientific workforce, it is imperative to increase participation of underrepresented minorities in the sciences (Malcom, 1996; Hrabowski, 2011). Broadening participation in the sciences requires us to reduce the barriers underrepresented minorities face in their undergraduate courses (Gibbs and Marsteller, 2016). Institutions of higher education trying to service more students and provide more flexibility are increasing course enrollments and using online technologies.

Two potential solutions for meeting the increasing demands on institutional spaces and resources are online and hybrid courses. In a 2012 survey, 33% of undergraduates had taken an online course (Allen and Seaman, 2014) and 55% were interested in taking one. Currently, there are not many online biology courses for biology majors (~10% of online biology courses in one study; Varty, 2016) and, thus, not much is known about the impact of online courses for majors. One of the few studies with biology majors was conducted at a community college and found that traditional-age students performed just as well in face-to-face versus online biology courses, but that performance of non-traditional age students was worse in online biology courses (Garman, 2012). Online biology courses for nonmajors are more common, and several studies found no impact of moving these courses online for student learning outcomes relative to a traditional lecture-based face-to-face course (Johnson, 2002; Hughes, 2008; Hauser, 2016). Thus, there is not a consensus in the literature about

Joel K. Abraham, *Monitoring Editor*

Submitted Aug 10, 2018; Revised Nov 19, 2018; Accepted Dec 17, 2018

CBE Life Sci Educ March 1, 2019 18:ar8

DOI:10.1187/cbe.18-08-0153

*Address correspondence to: Sat Gavassa (sgavassa@fiu.edu).

© 2019 S. Gavassa *et al.* CBE—Life Sciences Education © 2019 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/3.0>). “ASCB®” and “The American Society for Cell Biology®” are registered trademarks of The American Society for Cell Biology.

the merits of online courses versus traditional face-to-face courses for biology majors. Moreover, there are still questions about which student populations are best served by online courses and which are most affected.

Hybrid courses, in which instructor–student face time is reduced by half and replaced with online activities, also have the potential to reduce space and staffing needs and still allow direct student–instructor contact. Thus, hybrid courses are gaining attention as an alternative to traditional face-to-face classes or fully online courses. However, the literature reports conflicting results regarding the effect of hybrid courses on student learning. Adams and collaborators (2015) compared two sections of introductory microbiology; one section was made hybrid by replacing one of two weekly face-to-face lectures with an online lecture while keeping everything else equal. Student performance, as measured by exam scores, was lower in the hybrid section, especially for minority students. On the other hand, a hybrid introductory biology course with additional assignments showed learning gains for minority students and found no difference for nonminority students relative to a traditional face-to-face version of the course (Riffell and Merrill, 2005). Likewise, a hybrid section of introductory sociology with active learning was particularly effective in reducing the achievement gap for students of color compared with a traditional face-to-face lecture section (Luna and Winters, 2017). Thus, there are some hints in the literature that hybrid courses can have a positive effect on student learning relative to traditional face-to-face courses as long as class time is replaced with structured student-centered assignments rather than recorded passive lectures (Baeppler *et al.*, 2014; Crimmins and Midkiff, 2017). Studies comparing student performance in hybrids versus fully online courses are still lacking.

We hypothesize that a hybrid course with high levels of in-class engagement and active learning and highly structured to increase student interaction with the material could increase student performance, including performance of traditionally underrepresented students, while reducing in-class contact time. Active learning has been shown to improve student performance across many disciplines in science, technology, engineering, and mathematics in higher education (reviewed in Freeman *et al.*, 2014). Furthermore, student performance in introductory biology courses improves when active learning is combined with the addition of guiding course structures including: preclass assignments, in-class engagement activities, and postclass review assignments. This increased course structure reduces the achievement gap, even when sections grow larger in size and have higher ratios of students per instructor and/or graduate teaching assistants (Haak *et al.*, 2011; Eddy and Hogan, 2014).

We set out to study how a highly structured hybrid course that incorporated preparatory assignments, group work on problems in class, and review assignments impacted student performance relative to a traditional lecture and a fully online version of the same course. Our institution is a large Hispanic-serving research university with a large proportion of underrepresented and first-generation students, so we were particularly interested in the effect course format had for these groups. We found that students in the hybrid course performed better on exams compared with students in the lecture or fully online versions of the course. We present our course design and how it relates to student success. This work adds to the

emerging literature on hybrid courses and will inform curricular practices modeled after blended pedagogical strategies.

METHODS

We designed, taught, and assessed three versions of General Biology I: a traditional face-to-face, a fully online section, and a hybrid and high-structure course (50% online 50% face-to-face). Here, we report the results from the hybrid course in comparison with traditional face-to-face and fully online courses. The formats varied in contact time, with the online format having the least contact and face-to-face having the most. The three formats also varied in the number of structured assignments students were expected to engage in beyond exams. The hybrid format had the highest number of tasks, while the face-to-face format had the fewest number of graded items per week.

Description of the Course

General Biology I is an entry-level course with large enrollment (average annual course enrollment: 2303 ± 312 ; average section size: 367 ± 84). Enrollment for this course has grown by hundreds of students each year for the last few years. The course described in this study was offered at a large public Hispanic-serving institution in the Southeast as the first semester of a two-semester introductory biology sequence. This course is required for biology majors but is also taken by other science and engineering majors. The majority of the students in the course are in the pre-health track. Although multiple instructors teach this course in a year, all three sections used in this study were taught by the same instructor (S.G.) and covered the same topics. Course topics include general introduction to metabolism, molecular biology, genetics, and evolution. The face-to-face format was taught in Fall 2013, and the hybrid format was taught during Fall 2014. The online format was taught simultaneously with the other formats (Fall 2013 and Fall 2014). The formats differed in the number of hours students attended class in person and the number of required assignments (Table 1).

This course has an associated laboratory course, which is a required co-requisite. Labs are administered centrally for all sections of this course regardless of modality or instructor. All lab sections were fully in person, had the same curriculum, and were taught by teaching assistants, who are trained by the lab coordinator. There are between 30 and 40 lab sections per semester. Students enroll in the section that best fits their schedule.

Design of Face-to-Face, Hybrid, and Online Formats

Design of all three formats was informed by evidence-based practices from the discipline-based education research literature. Specifically, we will describe these courses in terms of the increased structure format (Freeman *et al.* 2011; Eddy and Hogan, 2014). A course with high structure refers to a course in which the instructor includes multiple activities as part of the course with the purpose of guiding student engagement with the course content. In a high-structure course, students: 1) complete preparation assignments before coming to class, 2) engage in active-learning exercises in class, and 3) complete weekly low-stakes review or practice assignments (Freeman *et al.*, 2011; Haak *et al.*, 2011). The documented advantages of a high-structure course include encouraging students 1) to

TABLE 1. Comparison of course formats showing that the hybrid format had the highest number of active-learning assignments, with activities due online similar to the online format, plus additional in-class activities, while the face-to-face format had the highest contact time per week, with students mostly passive in class, listening to the instructor lecture and occasionally answering iClicker questions

	Online	Hybrid	Face-to-face
In-person contact time: minutes per week	0	75	150
Course structure	Moderate	High	Moderate
Preclass LearnSmart assignments	Optional	Required	Optional
In-class iClicker questions per class period	NA	≤15%	≤15%
Peer discussions per class	Online discussion boards	≥60%	≤10%
Graded review assessments	Online quiz	Online quiz and in-class IF-AT	Online quiz
Time lecturing in class	NA	≤10%	≥80%
Time online for video lectures	75 minutes	75 minutes	NA

spend more time reading the textbook (Seaton *et al.*, 2014), 2) to come to class better prepared (Gross *et al.*, 2015), and 3) to attend class (Riffell and Sibley, 2004). Finally, students in high-structure courses have more opportunities to actively engage with the material, all of these advantages have been shown to improve student performance (Carini *et al.*, 2006; Eddy and Hogan, 2014; Freeman *et al.*, 2014).

Face-to-Face Format. The face-to-face course data included in this study are from one section of 229 students taught in three 50-minute class sessions per week, for a total contact time of 150 minutes per week. In the face-to-face format, lectures were delivered during class time using PowerPoints and occasional video clips. Each class session included 5–10 questions using the iClicker classroom response system. Students were asked to answer individually first. If only a small proportion of the students answered an iClicker question correctly, all students were asked to discuss the answer with their neighbors and the question was reopened. Students received 1 point for participating and 1 point for selecting the correct answer. The points for iClicker questions were extra credit for the face-to-face format; thus, not all the students used an iClicker. All three formats had a weekly online quiz. However, the lecture section had to coordinate the quiz questions with other instructors teaching other lecture sections that semester. The face-to-face format was moderately structured, as it had two of the three elements of a high-structure course, student in-class engagement activities and a weekly review assignment, but lacked preparatory assignments.

Hybrid Format. The hybrid course data are from four sections taught during Fall 2014, capped at 96 students each ($n = 94, 89, 90,$ and 93), that met once a week for a 75-minute class session. In the hybrid format, video lectures and other resources were delivered in an online self-paced course platform. On average, videos and out-of-class resources accounted for 75 minutes of online work. Face-to-face time was reduced to 75 minutes once a week (50% face-to-face). The main advantage of fewer meeting times is that a large section can be separated into multiple sections with reduced enrollment without using additional classroom space. Before class, students were expected to watch the assigned videos and read the textbook. To encourage students to complete their assigned work, we required students to complete a set of multiple-choice questions using the online

learning platform provided by the publisher. During the first half of class meetings, the instructor reviewed the main topics from the videos interspersed with iClicker questions (very similar to those used in the face-to-face format). During the second half of the class meeting, students took an individual assessment (~10 multiple-choice questions, ~10 minutes of class time). Later, for the remaining 30 minutes of class, students answered those same questions in small groups (3–6 students) using an immediate feedback assessment technique (IF-AT) that allowed them to immediately see which questions they answered correctly. While students worked in teams, they received guidance from the instructor and undergraduate peer learning assistants. The IF-AT is based on a five-choice question and was graded based on the number of attempts students took to arrive to the correct answer; if they answered correctly the first time, they received 5 points per question, 1 point was deducted for each failed attempt. Thus, if students took all five attempts to find the correct answer, they would still get 1 point, but if they gave up and did not find the correct answer, they received zero points. Finally, students took a review quiz once a week, as in all the other course formats. The hybrid sections were considered to be high-structure courses, as they contained all three elements: preclass preparatory assignments, in-class student engagement, and a review assignment.

Online Format. The online course data are from two sections, one taught in Fall 2013 and the other taught during Fall 2014 with 72 and 75 students, respectively (Figure 1). Students had the same video lectures and animations available online as the hybrid students, which amounted to ~75 minutes of video lectures per week. In addition, online students had a weekly assignment of six open-ended questions that they had to answer in at least one paragraph. Students could choose to complete the assignment individually or work with a team of six students (about half of students chose to work individually). Students were expected to watch the videos before answering the discussion questions, although there was no way to enforce this. Time on task answering the discussion questions was expected to be ~75 minutes as well. Finally, students took a review quiz at the end of the week. Online students only meet on campus for proctored exams twice in the semester (a midterm and a final exam). In the online format, a preparatory assignment was not required to be completed before students could begin the video lectures. However, students had to answer the six weekly

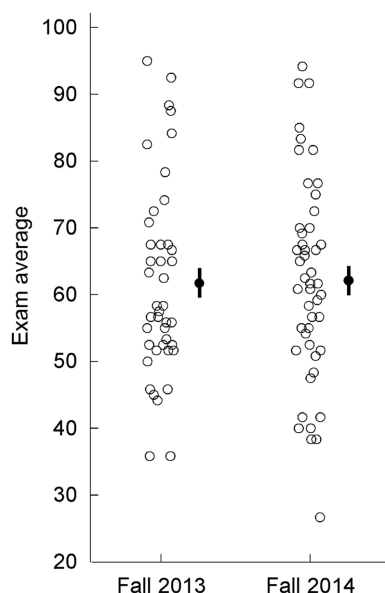


FIGURE 1. Exam scores for students in the online course format by semester. Open circles indicate individual data points; filled circles indicate mean; vertical lines indicate SE. One section of the online course was taught in the same semester as the face-to-face course, and another section was taught during the same semester as the hybrid course. We used a simpler version of the main regression model to test for differences in student performance in the online format between semesters (Exam score ~ SAT Math score + Semester). Our simple model significantly explained a portion of the variation in student performance ($R^2 = 0.129$, $F(2, 42) = 4.25$, $p = 0.021$). The SAT Math score predicted a significant portion of the variation in student performance (Estimate \pm SE: 0.057 ± 0.019 , $p = 0.0057$). However, the semester when the course was taught did not have a significant effect on student performance (Estimate \pm SE: 0.074 ± 3.49 , $p = 0.98$). Thus, we combined both online sections for our analyses.

discussion questions as they progressed through the lecture videos. We consider the discussion questions similar to in-class engagement. As with the hybrid sections, students completed the quiz at the end of the week as their postclass review assignment. Thus, the online format only contained two of the three elements of a high-structure course and is considered moderately structured.

Assessment of Student Learning

Here, we compare exam scores for the three course formats. Face-to-face and hybrid include the average of four proctored

exams, while the online format only had two proctored exams. Exams for face-to-face and hybrid format were identical and included the same number of chapters and the same number of questions (50 questions per exam, 200 questions total) and question prompts. Because the online course only included two proctored exams, each exam included twice as much material as the other two formats and was longer, with 60 questions per exam. However, those 120 questions (60 questions in two exams) were a subsample of the questions used in the hybrid and face-to-face exams (120 in online exams out of 200 questions in hybrid/face-to-face exams). The questions in these exams were all created by the instructor.

Exam Equivalence Analysis

Although the questions in the online exams were a subsample of the questions used in the hybrid and face-to-face exams, it is possible that the subsample had been skewed and did not reflect the level of difficulty of the other exams. Thus, to check that the subsample of questions selected for the online exam was equivalent to the difficulty level of the hybrid and face-to-face exams, we ranked all 200 questions using the Blooming Biology Tool by Crowe *et al.* (2008). We condensed the five Bloom's categories into three: low (knowledge and comprehension), medium (application), and higher order (analysis and synthesis). The questions were rated by two of the authors (R.B. and S.G.). The raters first ranked the questions independently. There was a 70.5% consensus in the independent ranking of the questions. The questions that had different rankings (29.5%), were discussed afterward, and we assigned a consensus rating. We did not find a significant difference in the proportion of high-, medium-, or lower-order questions when we compared the online exams with the hybrid/face-to-face exams, indicating that the subsample of questions included in the online exams was representative of the questions in the exams for the other formats ($\chi^2 = 2.272$, $df = 2$, $p = 0.32$). We found that 20.9% of the online exam questions and 23.5% of the questions in hybrid and face-to-face exams were either medium or higher order (Table 2). We consider this to be a high proportion compared with studies finding that exams for biology courses have on average only 6.7% of questions at medium or higher levels of Bloom's taxonomy (Momsen *et al.*, 2010).

Student Populations

Students were self-selected and were somewhat aware of the formats of the courses in which they were enrolling. Students enrolling for the online session had to pay an additional fee; thus, they were very aware that they were enrolling in a fully online class. On the other hand, while the hybrid course format included a note indicating the course was hybrid, most students

TABLE 2. Exam equivalence for hybrid and face-to-face exams, with online exam information shown in parentheses^a

	Total	High	Medium	Low
Number of questions	200 (120)	10 (2)	37 (23)	153 (94)
Proportion of questions	Hybrid/face-to-face (online)	5% (1.7%)	18.5% (19.2%)	76.5 (78.3%)

^aThe exams for the hybrid format and the face-to-face format were identical; we had four exams with 50 questions each (200 total questions). However, there were only two online exams with 60 questions each, which were a subsample of the questions used in the face-to-face/hybrid exams (120 out of 200 questions). There were no significant differences in the proportion of questions in each of these three categories for the face-to-face/hybrid exams and for the online exams—values indicated in parentheses ($\chi^2 = 2.272$, $df = 2$, $p = 0.32$). We condensed the five Bloom's categories into three: low (knowledge and comprehension), medium (application), and high (analysis and synthesis). Another study looking at multiple biology courses found that exams on average only include 6.7% of medium- or higher-order questions (Momsen *et al.*, 2010). Conversely, the exams shown here had between 20.9 and 23.5% (online and hybrid/face-to-face, respectively) of either medium- or higher-order questions.

TABLE 3. Student demographics in general biology course by course format showing the number of students included in the analyses for each category: sex, ethnicity/race, or college-level category in the online, lecture, and hybrid course formats (Figures 2 and 3)^a

	Online	Face-to-face	Hybrid
Total	74 (147)	124 (229)	212 (366)
Sex			
Male	31	56	93
Female	43	68	119
Ethnicity/race			
White	10	12	20
Hispanic	48	94	158
Black	16	18	34
College level			
Freshmen	11	67	138
Sophomores	21	28	41
Juniors	23	20	27
Seniors	19	9	6

^aThe number in parentheses indicates the total number of students enrolled in the course. The number of students included in the statistical analysis is smaller than the number of enrolled students, because we included only data from students who completed all the exams in each course. Thus, students who dropped a course or who missed at least one exam are not included in the analyses. Due to the low enrollment of some racial/ethnic groups, only white, Hispanic, and Black students were included in the analysis.

did not see it, based on how many students raised their hand when asked whether they were aware of the hybrid format during the first day of class. Moreover, online sections are capped at smaller numbers compared with hybrid and face-to-face sections. Because course registration opens in order of college level, senior and junior students are more likely to take the limited spots in the online sections (Table 3 and Figure 2). Furthermore, our analyses only included students who took all the exams on the course, as we calculated the average for all the exams. The number of students was also limited to those students for whom the school had records of their Scholastic Aptitude Test (SAT) scores. Students who dropped the course, missed an exam, or did not have an SAT score were not included in the analyses (410 students were included in the analyses out of 742 students enrolled).

Statistical Analyses

Differences in student demographics and college level (freshman, sophomore, etc.) across the course formats were tested using a chi-square nonparametric test. Because the study took place over 2 years, we were concerned about year-to-year variation. To increase our confidence that the results were due to course format and not to variation in student ability, we controlled for SAT Math score.

To determine whether course format was correlated with student performance on exams, we used a linear regression model with exam points as a continuous response variable. We included SAT Math scores, college level (freshman, sophomore, junior, or senior), gender (male or female), and race/ethnicity (Black, Hispanic, or white) as control variables. We used Akaike's information criterion (AIC) to compare multiple models including one or more interaction terms and selected the model that had the highest support (Table 4). The final model

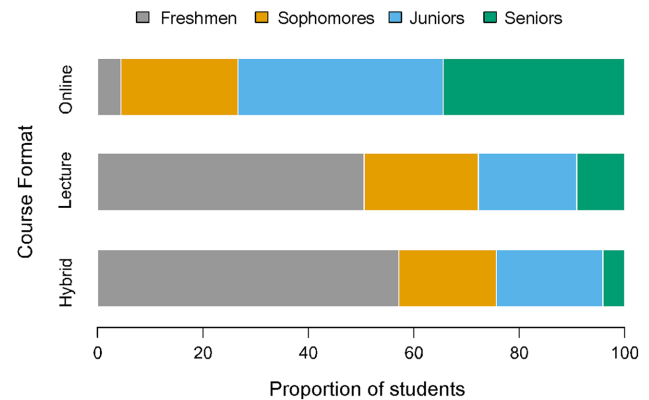


FIGURE 2. College level by course format. While freshmen make up the majority of the students in the hybrid and face-to-face formats (online vs. hybrid: $\chi^2 = 5.8713$, $df = 3$, $p = 0.118$), seniors and juniors make up the majority of the students in the online format ($\chi^2 = 115.43$, $df = 6$, $p < 0.001$). These differences are probably due to registration priority increasing with student seniority and limited availability of spots in the online sections.

predicted exam performance using SAT Math subscore/score, course format, college level, and race/ethnicity as control and predictive variables, as well as an interaction term between race/ethnicity and course format. We then looked for differences between the multiple student groups using Tukey's honestly significant difference (HSD) post hoc test.

Selected model: Outcome ~ SAT Math score + Course format + College level + Race/ethnicity + Race/ethnicity \times Course format

The online format was the only one taught during the two semesters of this study. We tested for differences in student performance between the two semesters when the online course was taught (Fall 2013 and Fall 2014) using a simple linear regression model with exam score as outcome and controlling for SAT Math score (Exam scores ~ SAT Math score + semester, online format only).

RESULTS

Model Validity

Our linear regression model, including SAT Math score, college level, ethnicity or race, and course format explained a significant proportion of the variation in student performance ($R^2 = 0.387$, $F(12, 397) = 20.87$, $p < 0.001$). The SAT Math score was a significant predictor of student performance for all the models in which it was included (all except the null model; $p < 0.0001$ for the selected model; Table 5). However, the AIC decreased as other factors were included in the model (Table 4).

Because the online format was taught during both semesters, we tested whether student performance in exams differed between semesters. Our smaller model including only the online format significantly explained a portion of the variation in student performance ($R^2 = 0.129$, $F(2, 42) = 4.25$, $p = 0.021$). As with the main model, SAT Math score predicted a significant portion of the variation in student performance ($p = 0.0057$). However, we did not find a significant effect of semester

TABLE 4. Comparison of linear regression models as predictors of exam performance with the top five models predicting exam performance ranked from highest support to lowest according to AIC and the highest weighted AICs (ω)

Model ^a	AIC	ω
SAT Math score + Course format + Race/ethnicity + Year level + Course format \times Race/ethnicity	3158.44	0.405
SAT Math score + Course format + Race/ethnicity + Year level + Sex \pm Course format \times Race/ethnicity	3159.44	0.246
SAT Math score + Course format + Race/ethnicity + Year level + Sex \pm Course format \times Race/ethnicity + Course format \times SAT Math score	3159.92	0.193
SAT Math score + Course format + Race/ethnicity + Year level + Sex \pm Course format \times Race/ethnicity + Course format \times Year level	3160.58	0.139
SAT Math score + Course format + Race/ethnicity + Year level + Sex \pm Course format \times Race/ethnicity + Course format \times Year level	3165.73	0.011
SAT Math score + Course format + Race/ethnicity + Sex	3166.82	0.006
SAT Math score	3297.06	0
Null	3334.99	0

^aBold type indicates additional interaction terms included in the different models.

when the course was taught on student performance ($p = 0.98$; Figure 1). Thus, we combined the data from both online semesters in our analyses.

Effect of Sex

Sex data were obtained from student self-reported selection on a binary choice when completing the school application. All course formats had a predominantly female enrollment (52% in online, 53% in hybrid, and 58% in face-to-face; Table 3); there were no significant differences between the course formats in the proportion of female students ($\chi^2 = 1.4604$, $df = 2$, $p = 0.4818$). Moreover, after accounting for SAT Math subscore, our model did not reveal any difference in performance between male and female students. In fact, the strength of the model did not improve when student sex was included. Comparing the weights of the models, the model without sex as a predictor is 1.6 times more likely to be the best model in terms of Kullback–Leibler discrepancy than the model including sex (Wagenmakers and Farrell, 2004). Thus, sex was not included in the selected model (Table 4).

Effect of College Level

The college level of students enrolled in each course format was different between online and the other two formats (hybrid and face-to-face). The online format had a majority of seniors and juniors, while hybrid and face-to-face had a majority of freshman students ($\chi^2 = 115.43$, $df = 6$, $p < 0.001$; Table 3). There were no significant differences in the college level between students enrolled in the face-to-face and the hybrid course formats (face-to-face vs. hybrid: $\chi^2 = 5.8713$, $df = 3$, $p = 0.118$). After accounting for Math SAT subscore, college level was a strong predictor of student performance (Table 5). Seniors and juniors scored ~6 points higher on exams compared with freshmen and sophomores. Post hoc multiple-comparison analyses revealed no differences between freshman and sophomore exam scores or between junior and senior exam scores but did show that juniors and seniors outperformed freshmen and sophomores.

Effect of Race or Ethnicity and Course Format

Not surprisingly, given that we are a Hispanic-serving institution, all course formats also had a majority of Hispanic students

(Figure 3). However, there were no significant differences in the ethnic/racial composition of students across the course formats ($\chi^2 = 4.987$, $df = 8$, $p = 0.759$; Table 3). We used Hispanic students as our reference group in our analyses. After we controlled for SAT Math subscore, college level, and sex, Hispanic students in the hybrid course format scored 26% higher and 12% higher than Hispanic students in the face-to-face format and online format, respectively (Table 5). On the other hand, we found that, even after we controlled for SAT Math subscore,

TABLE 5. Linear regression model with coefficients of the regression model represented in terms of raw points \pm SE out of 100 points^a

Regression coefficients	Estimate \pm SE	<i>p</i> value
Model intercept	37.03 \pm 4.07	
Student achievement		
SAT Math score	0.054 \pm 0.008	<0.001
College level (Reference level: Freshmen)		
Senior	8.73 \pm 2.26	0.003
Junior	4.81 \pm 1.62	0.002
Sophomore	-0.25 \pm 1.46	0.862
Race/ethnicity (Reference level: Hispanic)		
Black	-4.91 \pm 2.13	0.021
White	-0.96 \pm 2.66	0.72
Course format (Reference level: Hybrid)		
Traditional lecture	-14.85 \pm 1.46	<0.001
Online	-9.53 \pm 1.95	<0.001
Course format*Race/ethnicity (Reference level: Hybrid*Hispanic)		
Lecture*Black	0.45 \pm 3.59	0.901
Online*Black	-3.87 \pm 3.94	0.325
Lecture*white	-1.10 \pm 4.37	0.817
Online*white	15.15 \pm 4.78	0.002

^aThe categorical variable “college level” represents achievement by college level relative to the achievement of students at the freshman college level for students in all three course formats. The sex category represents the gains of male students in all three course formats relative to female students. The race/ethnicity category represents the performance by racial group relative to the performance of Hispanic students. After we controlled for college level, sex, and race/ethnicity, the coefficients of the regression model indicate that students in the hybrid format outperform students in either online or face-to-face format. The complete model explains more than a third of the variation in student achievement ($R^2 = 0.387$, $F(12, 397) = 20.87$, $p < 0.001$).

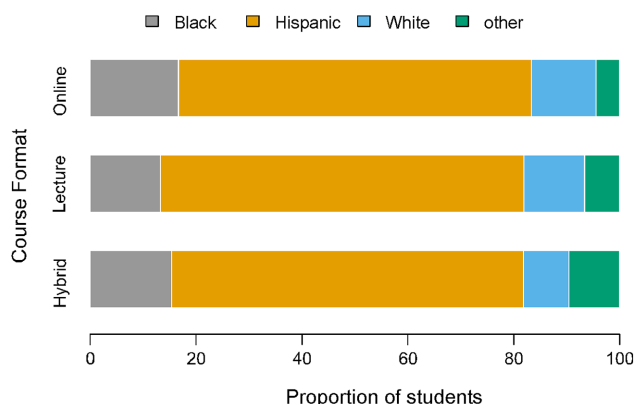


FIGURE 3. Ethnicity or race by course format. At our Hispanic-serving institution, the majority of our students self-identified as Hispanic. The next most common category included students who identified themselves as Black, although this does not distinguish between African American, Haitian, Black Caribbean, and other categories. The next category included students who identified as white. A very small number of students identified themselves as Native American, Alaskan, Native Hawaiian/Pacific Islander, or Caribbean, and were grouped into the “other” category.

Black students scored lower on exams compared with other student groups ($p = 0.024$; Table 5 and Figure 4). There were no differences in exam performance by race in either the hybrid or face-to-face formats (Table 5 and Figure 4). However, performance was lowest across all racial/ethnic groups in the face-to-face format, while performance was highest for all racial/ethnic groups in the hybrid format. Only the online format had significant differences in student performance by race. White students in the online format outperformed Hispanic students by almost 15 points (22%) and Black students by 19 points (28%). In summary, Black and Hispanic students had their highest scores in the hybrid format. Conversely, white students had their highest scores in the online and hybrid formats (Table 5 and Figure 4).

DISCUSSION

We set out to study how student performance would change moving from a completely face-to-face introductory biology course to hybrid and fully online formats. The lecture and online courses had similar levels of structured activities, while the hybrid format allowed for the highest amount of structure. We found that, after we controlled for prior performance (Math SAT subscore), the hybrid format resulted in the highest performance across all ethnic/racial groups. The results of Tukey’s HSD show that the lowest scores were found in the face-to-face format. In addition, performance gaps between Hispanic, Black, and white students were mostly found in the online format. There was a persistent effect of college level. As shown in other studies, we found that juniors and seniors scored higher than freshmen and sophomores (Adams *et al.*, 2015, 2016).

Effect of Race/Ethnicity and Course Format

As minority-serving institution with a majority of Hispanic students (61% of the total student population), we set Hispanic students as the reference group for our analyses. The model

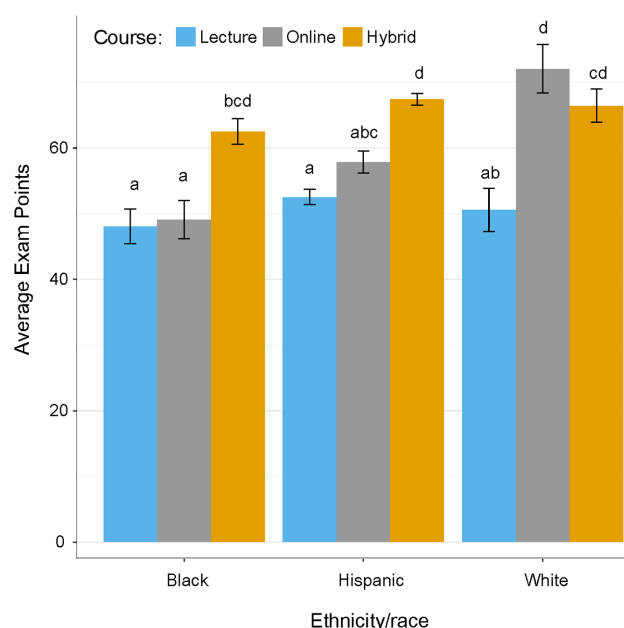


FIGURE 4. Interaction between course format and race/ethnicity. Means and SEs as predicted by our linear model ($R^2 = 0.387$, $F(12, 397) = 20.87$, $p < 0.001$), broken down by course format and race/ethnicity; both factors had significant impact on student performance (Table 4). The letters indicate significantly different groups based on Tukey’s HSD post hoc test. For Black and Hispanic students, highest performance was observed in the hybrid format. Moreover, their performance is undistinguishable from that of white students, showing no differences by race in the hybrid format. On the other hand, the lowest performance for all students, regardless of race/ethnicity, was in the face-to-face format. In the online format, white students performed the highest.

shows a performance gap between Black and Hispanic students overall (Table 5). Hispanic students differed from white students only in the online format, where white students had the highest performance (Figure 4). Another study in an introductory biology course found that minority students tend to underperform in face-to-face and hybrid formats (Adams *et al.*, 2015). However, a study in an upper-division course found no significant differences in performance between underrepresented and traditional students (Adams *et al.*, 2016). Conversely, we found that Black and Hispanic students had their respective highest performance in the hybrid format compared with students of the same race/ethnicity in online or face-to-face formats.

For all three formats, the exam questions included a large number of higher-order application- and integration-type questions (Table 2), which have been shown to reduce the performance of students from traditionally underrepresented groups (Wright *et al.*, 2016). While the hybrid and face-to-face courses had identical exams, the questions in exams for the online course were a subsample of the hybrid and face-to-face exams. Our exam equivalency analysis showed a similar proportion of questions for each level of difficulty in the subsample for online exams. Although we did not look at scores for individual exam questions, we found that, in the hybrid course, there was no significant difference in the average score for all exams in the

course comparing white and Hispanic students, and the difference was greatly reduced for Black students. The increased course structure and incorporation of active-learning strategies through the flipped-classroom approach were factors that likely influenced these results. Likewise, Balaban and collaborators (2016) found that a flipped class in economics increased student performance, especially in higher-order learning, such as application level. Moreover, they found higher effects for low-income students (Balaban *et al.*, 2016).

It is possible that, because the hybrid course halved the time students needed to be present on campus, the hybrid format by itself could be beneficial to students by giving them more flexibility in their schedules and thus reducing conflicts between their academic and personal responsibilities. However, existing evidence suggests that making a course hybrid per se does not seem to improve student performance, and instead it may decrease learning. Adams *et al.* (2015) found that, for a microbiology course, sophomore students had lower performance in a hybrid course than students in a traditional face-to-face format. However, their hybrid course had the same amount of time spent in active-learning activities as their face-to-face section. The only difference in their course formats was that, in the hybrid format, they delivered one lecture per week face-to-face and the other lecture was streamed online (Adams *et al.*, 2015). Thus, schedule flexibility alone does not seem to explain learning gains in the hybrid format. It is important to remember that the hybrid course we describe here is not just hybrid (50% face-to-face and 50% online), but is also a highly structured course with preclass preparatory assignments, in-class student engagement, and postclass review assignments (Freeman *et al.*, 2011; Haak *et al.*, 2011). Thus, it is possible that schedule flexibility adds to the beneficial effects in student performance when combined with a highly structured student-centered design.

In our study, the three course formats differed in both contact time and structure. While the online course had the least contact time and the face-to-face format had the greatest contact time, the hybrid format had the highest structure, whereas the other two formats were moderately structured (as defined in Haak *et al.*, 2011). Thus, the positive results of the hybrid format may be a combination of high structure and moderate contact time.

Effect of In-Class Meetings

While Black and Hispanic students had their highest scores in the hybrid format, white students had their highest scores in the online format. Although the online format was not more highly structured than the lecture course, as strictly defined by Haak and collaborators (2011), the online format had more opportunities to engage with the material in the form of weekly assignments, such as team discussions. The additional interaction with the material seems to have favored white students in the online environment in comparison with lecture format. Hispanic students also performed better in the online format than in the lecture course, although not as well as they did in the hybrid format. Conversely, there was no significant difference for Black students when comparing online and lecture, only the hybrid format showed an improvement in their performance (Figure 4).

Although the hybrid and online courses had identical online resources, the hybrid had additional in-class active-learning

exercises done in teams and with support from undergraduate learning assistants and the instructor, simultaneously increasing the contact time and the structure. It is possible that physical contact with the instructor and with other students had a higher impact for Black and Hispanic students than for white students. Some studies have found that group learning is particularly beneficial for underrepresented students (Peters, 2005). Interaction with other students, such as discussing the course content, has been shown to increase the sense of belonging and improve performance of students from Hispanic backgrounds compared with students from Caucasian backgrounds (Hurtado and Carter, 1997; Savani *et al.*, 2013). Furthermore, aspects such as the structure of the teamwork and the composition of the teams have an effect in the learning gains (Jensen and Lawson, 2011).

Even though the face-to-face format had the highest contact time, it also resulted in the lowest performance. Although the face-to-face format satisfies the condition of in-class engagement, the quality of the in-class engagement is very different from that of the hybrid format. While students in the face-to-face format spent more time seated in the same room, they did not actually spend more time interacting. When students interacted in the face-to-face format, those interactions were informal and unstructured. Students were only asked to discuss with their neighbors after a clicker question returned a majority of incorrect responses. Even then, the interaction was voluntary, unstructured, and unmonitored. Observations of student discussions have found that a large proportion of the conversations students have when discussing a clicker question in an unstructured manner are unproductive and even misleading (James and Willoughby, 2011). Conversely, in the hybrid format, students worked in permanent teams on the IF-AT assignment, which was graded, and received guidance from the peer learning assistants and the instructor.

Limitations of the Study

The hybrid and the face-to-face formats were taught in different semesters, whereas the online course was offered in both semesters. However, the online course was taught concurrently with the other two formats, and we did not find any significant differences in student performance in this course from one semester to the next (Figure 1). We cannot rule out that the changes in student performance can be the result of year-to-year variation. However, we did not find evidence of student performance being affected by year-to-year variation when comparing the online sections.

Students self-enrolled into each of these formats, and as noted by the differences in college level between online students and those in the other two formats, we see that particular student groups enroll preferentially in a given format. To control for these differences, we included several of these student variables in the model for our analyses by accounting for SAT Math subscores, college level, and race/ethnicity.

We obtained student ethnicity/racial data from student self-reported designations in their school applications. We disaggregated underrepresented students into Black and Hispanic groups, as these groups are subject to different sociological factors affecting their performance (Carpenter *et al.*, 2006). However, our categorization still clusters multiple nationalities and

ethnic groups together. For example, our school has a large proportion of Haitian students, some of whom recently immigrated to the United States, who are clustered with African-American students. Similarly, our Hispanic student population is very diverse, differing, for example, in exposure to college culture and degree of English proficiency. This diversity of backgrounds is not captured in the racial/ethnic categories included in this study.

CONCLUSIONS

Increased course structure and active learning have been shown to differentially benefit underrepresented students (Haak *et al.*, 2011; Eddy and Hogan, 2014). Our study supports the hypothesis that increased course structure improves the performance of traditionally underrepresented students, even with a reduction in face-to-face instruction. The moderate structure of the online course was sufficient to increase the performance of white students, whose performance was indistinguishable in online and hybrid formats. However, the increased structure of the course design, as well as the increased guidance of the peer face-to-face discussions, greatly favored the performance of Black and Hispanic students in the hybrid format. Nonetheless, this study leaves open the question as to whether it was the peer-peer and student-instructor engagement or the increased structure that resulted in increased performance of the hybrid format. This study adds to the literature supporting positive results of hybrid courses on students' performance when accompanied by an increase in course structure with more active learning and student-centered instruction.

ACKNOWLEDGMENTS

We thank the editor and two anonymous reviewers for insightful comments and suggestions. This project was approved by the Florida International University IRB (IRB-15-0136). This work was partially supported by Howard Hughes Medical Institute grant no. HHMI 52006924 and grant no. HHMI 52008097.

REFERENCES

- Adams, A. E. M., Garcia, J., & Traustadóttir, T. (2016). A quasi experiment to determine the effectiveness of a "partially flipped" versus "fully flipped" undergraduate class in genetics and evolution. *CBE—Life Sciences Education*, 15(2), ar11. <https://doi.org/10.1187/cbe.15-07-0157>
- Adams, A. E. M., Randall, S., & Traustadóttir, T. (2015). A tale of two sections: An experiment to compare the effectiveness of a hybrid versus a traditional lecture format in introductory microbiology. *CBE—Life Sciences Education*, 14(1), ar6. <https://doi.org/10.1187/cbe.14-08-0118>
- Allen, I. E., & Seaman, J. (2014). *Grade change: Tracking online education in the United States* (11th annual report). Babson Survey Research Group and Quahog Research Group. www.onlinelearningsurvey.com/reports/gradechange.pdf
- Baepler, P., Walker, J. D., & Driessen, M. (2014). It's not about seat time: Blending, flipping, and efficiency in active learning classrooms. *Computers & Education*, 78, 227–236. <https://doi.org/10.1016/j.compedu.2014.06.006>
- Balaban, R. A., Gilleskie, D. B., & Tran, U. (2016). A quantitative evaluation of the flipped classroom in a large lecture principles of economics course. *Journal of Economic Education*, 47(4), 269–287. <https://doi.org/10.1080/J00220485.2016.1213679>
- Carini, R. M., Kuh, G. D., & Klein, S. P. (2006). Student engagement and student learning: Testing the linkages. *Research in Higher Education*, 47(1), 1–32. <https://doi.org/10.1007/s11162-005-8150-9>
- Carpenter, D. M., Ramirez, A., & Severn, L. (2006). Gap or gaps: Challenging the singular definition of the achievement gap. *Education and Urban Society*, 39(1), 113–127. <https://doi.org/10.1177/0013124506291792>
- Crimmins, M. T., & Midkiff, B. (2017). High structure active learning pedagogy for the teaching of organic chemistry: Assessing the impact on academic outcomes. *Journal of Chemical Education*, 94(4), 429–438. <https://doi.org/10.1021/acs.jchemed.6b00663>
- Crowe, A., Dirks, C., & Wenderoth, M. P. (2008). Biology in Bloom: Implementing Bloom's taxonomy to enhance student learning in biology. *CBE—Life Sciences Education*, 7(4):368–381. <https://doi.org/10.1187/cbe.08-05-0024>
- Eddy, S. L., & Hogan, K. A. (2014). Getting under the hood: How and for whom does increasing course structure work? *CBE—Life Sciences Education*, 13(3), 453–468. <https://doi.org/10.1187/cbe.14-03-0050>
- Freeman, S., Eddy, S. L., & McDonough, M. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences USA*, 111(23), 8410–8415. <http://doi.org/10.1073/pnas.1319030111>
- Freeman, S., Haak, D., & Wenderoth, M. P. (2011). Increased course structure improves performance in introductory biology. *CBE—Life Sciences Education*, 10(2), 175–186. <https://doi.org/10.1187/cbe.10-08-0105>
- Garman, D. E. (2012). *Student success in face-to-face and online sections of biology courses at a community college in East Tennessee*. Johnson City: East Tennessee University.
- Gibbs, K. D., & Marsteller, P. (2016). Broadening participation in the life sciences: Current landscape and future directions. *CBE—Life Sciences Education*, 15(3), ed1. <https://doi.org/10.1187/cbe.16-06-0198>
- Gross, D., Pietri, E. S., Anderson, G., Moyano-Camihort, K., & Graham, M. J. (2015). Increased preclass preparation underlies student outcome improvement in the flipped classroom. *CBE—Life Sciences Education*, 14(4), ar36. <https://doi.org/10.1187/cbe.15-02-0040>
- Haak, D. C., HilleRisLambers, J., Pitre, E., & Freeman, S. (2011). Increased structure and active learning reduce the achievement gap in introductory biology. *Science*, 332(6034), 1213–1216. <https://doi.org/10.1126/science.1204820>
- Hauser, L. (2016). An examination of the predictive relationship between mode of instruction and student success in introductory biology. *Inquiry: The Journal of the Virginia Community Colleges*, 20(1), 49–60.
- Hrabowski, F. A. (2011). Boosting minorities in science. *Science*, 331(6014), 125–125. <https://doi.org/10.1126/science.1202388>
- Hughes, L. (2008). Construction and evaluation of an online microbiology course for nonscience majors. *Journal of Microbiology & Biology Education*, 9(1), 30–37. www.ncbi.nlm.nih.gov/pmc/articles/PMC3577150/
- Hurtado, S., & Carter, D. F. (1997). Effects of college transition and perceptions of the campus racial climate on Latino college students' sense of belonging. *Sociology of Education*, 70(4), 324. <https://doi.org/10.2307/2673270>
- James, M. C., & Willoughby, S. (2011). Listening to student conversations during clicker questions: What you have not heard might surprise you! *American Journal of Physics*, 79(1), 123–132. <https://doi.org/10.1119/1.3488097>
- Jensen, J. L., & Lawson, A. (2011). Effects of collaborative group composition and inquiry instruction on reasoning gains and achievement in undergraduate biology. *CBE—Life Sciences Education*, 10(1), 64–73. <https://doi.org/10.1187/cbe.10-07-0089>
- Johnson, M. (2002). Introductory biology online. *Journal of College Science Teaching*, 31(5), 312–317.
- Luna, Y. M., & Winters, S. A. (2017). "Why did you blend my learning?" A comparison of student success in lecture and blended learning introduction to sociology courses. *Teaching Sociology*, 45(2), 116–130. <https://doi.org/10.1177/0092055X16685373>
- Malcom, S. M. (1996). Science and diversity: A compelling national interest. *Science*, 271(5257), 1817. <https://doi.org/10.1126/science.271.5257.1817>
- Momsen, J. L., Long, T. M., Wyse, S. A., & Ebert-May, D. (2010). Just the facts? Introductory undergraduate biology courses focus on low-level cognitive skills. *CBE—Life Sciences Education*, 9(4), 435–440. <https://doi.org/10.1187/cbe.10-01-0001>
- Peters, A. W. (2005). Teaching biochemistry at a minority-serving institution: An evaluation of the role of collaborative learning as a tool for science mastery. *Journal of Chemical Education*, 82(4), 571. <https://doi.org/10.1021/ed082p571>

- Riffell, S. K., & Merrill, J. (2005). Do hybrid lecture formats influence laboratory performance in large, pre-professional biology courses? *Journal of Natural Resources and Life Sciences Education*, 34, 96–100.
- Riffell, S. K., & Sibley, D. F. (2004). Can hybrid course formats increase attendance in undergraduate environmental science courses? *Journal of Natural Resources and Life Sciences Education*, 33, 16–20. Retrieved June 30, 2015, from www.lon-capa.org/papers/E03-16.pdf
- Savani, K., Alvarez, A., Mesquita, B., & Markus, H. R. (2013). Feeling close and doing well: The prevalence and motivational effects of interpersonally engaging emotions in Mexican and European American cultural contexts. *International Journal of Psychology*, 48(4), 682–694. <https://doi.org/10.1080/00207594.2012.688131>
- Seaton, D. T., Kortemeyer, G., Bergner, Y., Rayyan, S., & Pritchard, D. E. (2014). Analyzing the impact of course structure on electronic textbook use in blended introductory physics courses. *American Journal of Physics*, 82(12), 1186–1197. <https://doi.org/10.1119/1.4901189>
- Varty, A. K. (2016). Options for online undergraduate courses in biology at American colleges and universities. *CBE—Life Sciences Education*, 15(4), ar58. <https://doi.org/10.1187/cbe.16-01-0075>
- Wagenmakers, E. J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin and Review*, 11(1), 192–196. <https://doi.org/10.3758/BF03206482>
- Wright, C. D., Eddy, S. L., Wenderoth, M. P., Abshire, E., Blankenbiller, M., & Brownell, S. E. (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>