A Remote Instructor Like Me: Student–Teacher Congruence in Online, High School Courses

Jennifer Darling-Aduana iD

Georgia State University

Students belonging to marginalized groups experience positive impacts when taught by a teacher of the same race, ethnicity, and gender. The unique nature of standardized, asynchronous online course taking allows for greater separation of any possible educational benefits of student versus teacher-driven mechanisms contributing to these improved outcomes. Using a student-by-course fixed effect strategy on data from a large urban school district, I examined associations between whether students experienced racial/ethnic or gender congruence with their remote instructor and both engagement and learning outcomes. Students who identified as Black demonstrated higher rates of engagement, although no difference in achievement, within lessons taught by a same-race remote instructor. I find that representation is associated with engagement even when instructors follow closely scripted lessons, representation occurs in only small doses, and instruction occurs in an impersonal setting.

Keywords: *student–teacher match, representation, online learning*

RESEARCHERS have established positive impacts for students belonging to marginalized groups taught by teachers of the same race/ethnicity and gender (Dee, 2005; Egalite et al., 2015; Gershenson et al., 2016; Gershenson et al., 2017; Grissom et al., 2015; Keiser et al., 2002; Xu & Li, 2018). Sometimes these positive impacts are teacher-based when sharing demographic characteristics results in an increased willingness by teachers to make decisions favorable to in-group members (Nicholson-Crotty et al., 2011). Student-based mechanisms may also produce more positive outcomes due to differences in how students interpret their teachers' actions (Grissom et al., 2015). For instance, students taught by a teacher of the same race/ethnicity or gender might be less likely to experience stereotype threat—when a negative stereotype about their identity is activated due to concern that they might confirm that negative stereotype (Steele, 1997; Steele & Aronson, 1995). Relatedly, being taught by an instructor of the same identity might activate a role model effect, whereby instructors with similar ascriptive characteristics act "as behavioral models, representing the possible, and being inspirational" as a means to improve student motivation, self-efficacy, and subsequent performance (Morgenroth et al., 2015, pp. 465).

Understanding the mechanisms behind these "match effects" is important because student–teacher congruence appears to alleviate systemic inequities in a manner that improves the educational experiences and outcomes of student disproportionately harmed by these systems without removing resources or otherwise reducing the quality of educational experiences for students from nonmarginalized (or less marginalized) groups. In other words, there is an element to representation in the classroom that has yet to be fully identified that has the potential to pinpoint long withstanding bias and disparities in how children are educated in the United States. Beyond replicating the positive effects associated with student−teacher congruence by increasing representation in the classroom, the knowledge gained through probing the relative strength and efficacy of likely mechanisms can be leveraged to develop new systems and interventions to mitigate educational opportunity gaps.

Based on evidence of positive associations between student–teacher congruence and academic outcomes for students belonging to marginalized groups, researchers have proposed policy interventions that aim to increase the diversity of the teacher workforce, integrate culturally responsive teaching practices, and minimize teacher implicit bias with the goal of improving the academic experiences of students belonging to historically marginalized groups (Dee, 2005; Egalite et al., 2015; Gershenson et al., 2016, 2017; Grissom et al., 2015; Irizarry, 2015; Joshi et al., 2018; Lindsay & Hart, 2017). However, research has yet to establish the extent to which the mechanisms that support improved academic outcomes translate to an online instructional setting. Furthermore, this study employs a novel analytic strategy made possible through the detailed session-level data collected by the online course system studied. I leveraged these data to contribute to the nascent research base on which mechanisms through which student–teacher congruence might operate should be targeted for policy interventions

(i.e., Baker et al., 2018; Chin et al., 2020). For these reasons, I examined the following research questions.

Research Question: 1: What is the distribution of remote instructor identities in the 30 most enrolled in online courses in a large, urban district?

Research Question 2: To what extent is being taught by an instructor of the same race/ethnicity or gender associated with course outcomes when educational content is delivered through a standardized, prerecorded video?

I accomplished this by examining variation in student engagement and achievement patterns across different lessons when students were exposed to remote instructors of different racial/ethnic and gender identities within the same course. Because educational content was delivered through a prerecorded video that was scripted by an instructional design team, the instructor could not adjust the content, their instructional delivery, or interactions with each student in the asynchronous, online instructional environment studied. A study of spaces for teacher discretion in courses developed by the same vendor examined here found that there was no detectable deviation from the script provided by the company's curriculum development team in 40% of lessons (Darling-Aduana & Hemingway, 2021). Personalization of lesson content or delivery in the other 60% of lessons was limited to small modifications such as greeting students at the beginning of the class, offering encouragement, or emphasizing certain words in the script to convey humor (Darling-Aduana & Hemingway, 2021). Furthermore, the type of hybrid blended model implemented, where students completed at least part of their online course within a school setting is one of the most common realizations of technology-based learning, as it is easily integrated within a traditional schooling framework (Christensen et al., 2013). Leveraging data collected from this setting allowed for a nuanced examination of whether students modified their behavior in response to racial/ethnic or gender congruence with their instructor in a setting where teacher-based representation mechanisms were not feasible.

However, as I was unable to remove all potential sources of bias from my estimates, results should be interpreted as descriptive and not causal. Furthermore, this analysis assumes that exposure to an instructor for a single lesson and within a fully asynchronous, prerecorded online course environment is sufficient for representation effects to occur. Last, it is important to note that the student sample studied consisted primarily of students enrolled in online courses for the purpose of credit recovery. While an important subgroup to examine, this caveat should be noted when attempting to generalize findings to a wider population.

Prior Literature on Student−Teacher Congruence

In Dee's (2004) seminal study, White and Black elementary students of both genders achieved 3% to 6% higher reading scores when randomly assigned to a same-race teacher, with larger benefits observed among students who received Free or Reduced-Price Lunch (FRL). Using the National Education Longitudinal Study of 1988, Dee (2005) demonstrated that the benefits of race and gender congruence generalized to the national context. More recent studies confirm this phenomenon using Florida and Tennessee administrative data (Egalite et al., 2015; Joshi et al., 2018). The benefits of race congruence persist to high school graduation and college attendance (Gershenson et al., 2017). Beyond educational attainment, same-race teachers are also less likely to discipline or refer students from minoritized backgrounds to special education services and more likely to refer them to advanced courses (Grissom et al., 2015; Lindsay & Hart, 2017; Wright et al., 2017).

Researchers have identified similar effects among students who experienced gender congruence with their teachers in STEM (science, technology, engineering, and mathematics) fields (Keiser et al., 2002; Xu & Li, 2018). Student–teacher congruence is particularly salient in these subjects for female students, as society imposes norms of math as male-dominated increasingly as children progress through elementary into secondary school (Keiser et al., 2002; Solanki & Xu, 2018; Xu & Li, 2018). Keiser et al. (2002) established that eighth-grade students who identified as female achieved higher math scores when taught by a female math teacher. Similarly, Xu and Li (2018) observed higher female math performance and self-perceived math ability when they randomly assigned junior high students to female teachers. These positive effects are due at least in part to a role model effect, which results in reduced identity incongruence for female students when taught by a female teacher (Solanki & Xu, 2018).

Representation Within Online Courses

With a single exception (Baker et al., 2018), studies examining representation in schools focus on traditional, face-to-face classrooms. Prior to COVID-19, approximately 14% of secondary school students enrolled in online courses nationally (Gemin et al., 2015), with many more students now receiving at least a portion of their instruction virtually. As such, continued study of the online instructional environment is merited. Due to the dearth of information on instruction and instructors within the predominately for-profit online course marketplace, documenting the demographic characteristics of instructors hired by one of the largest course vendors in the United States represents a contribution.

More specifically, as online-based courses change the role of teachers in the classroom, the effects of representation may be implemented and realized differently. For

instance, it is possible that because students know the instructor cannot view or respond to them that stereotype threat might not be triggered in the online context (Steele, 1997; Steele & Aronson, 1995). Alternatively, if students appear to benefit, providing prerecorded video lectures delivered by instructors belonging to minoritized groups could be a stop-gap measure that could be implemented more quickly and cost effectively while long-term solutions (such as diversifying the teacher workforce) are being implemented. Knowing whether, and under what circumstances, online learning could be used to minimize bias in this manner has important equity implications.

Furthermore, the standardized, prerecorded nature of the online course system studied allowed for the isolation of potential confounding influences. Not only does the online course platform remove the possibility of active representation mechanisms requiring teacher discretion but it also limited the enactment of some potential passive mechanisms (i.e., some forms of role model effect that require the development of a student–teacher relationship).

However, to understand the potential mechanisms underlying representation in online classrooms, it is first necessary to understand what student–teacher interactions look like in these contexts. Most online courses, including those examined in this study, tend toward reductionism, or the packaging of easily digestible facts into lecture-based instruction with frequent multiple-choice assessments (Darling-Aduana, 2021; Heinrich et al., 2019; Herrington et al., 2006; McLoughlin & Oliver, 2000). The use of a common lecture video using a single script means that the only variation in instructional content and delivery emerges from how students interpret and respond to variations in the remote instructor delivering content and the lesson-specific content itself. It was this standardized curricular content and instructional delivery that I leveraged to better understand the underlying mechanisms through which representation operates within the asynchronous online course structure studied.

Disentangling Potential Mechanisms

There is a wide literature base establishing that teachers' discretionary behaviors likely contribute to positive associations between student outcomes and race/ethnicity and gender congruence. Examples include evidence of teacher implicit bias (Baker et al., 2018; Chin et al., 2020; Gershenson et al., 2016; Irizarry, 2015) as well as differential rates of responsiveness (Baker et al., 2018; Grissom et al., 2015), disciplinary action (Fish, 2017; Grissom et al., 2009; Grissom et al., 2015; Lindsay & Hart, 2017), and gifted referrals (Fish, 2017; Grissom et al., 2015).

What is less clear, and what this study aims to identify, is the extent to which student-based mechanisms might also explain positive associations with student–teacher congruence. Student-based mechanisms include changes in how students perceive their teachers and themselves when taught by a demographically congruent teacher (Grissom et al., 2015). For instance, students might be less likely to experience stereotype threat when taught by a teacher of the same race/ethnicity or gender (Dee, 2005; Steele, 1997; Steele & Aronson, 1995; Xu & Li, 2018). Same-race/ethnicity and gender teachers also might serve as positive role models to students with similar identities or whose interactional styles are based in a similar cultural background, who subsequently engage more in educational tasks (Kao, 2004; Xu & Li, 2018). Relatedly, there is evidence that teachers belonging to minoritized groups are more likely to induce positive behaviors in their students, termed coproduction inducement (Lim, 2006; Lindsay & Hart, 2017).

In an experiment that randomly assigned student names with specific racial and gender identities to researcher-generated discussion board comments within a postsecondary, online course, Baker et al. (2018) found that instructors were more responsive to students with names connoting White, male identities. When students were exposed to the same fictional peer discussion posts with randomly assigned racial and gender identities, the researchers identified no evidence of differential student responsiveness to peers based on those same connoted identities apart from White, female students, who were more likely to engage with students with the same supposed identities (Baker et al., 2018). This study provides evidence that teachers' discretionary actions might be more influenced by the assumed racial and gender identity of a student than students in peer-to-peer interactions, at least in the virtual, postsecondary classroom environment studied. However, this finding does not preclude the possibility that the racial and gender identity of the teacher might influence student behavior in a way that peer identity did not.

Continued research in this vein is needed to continue disentangling possible mechanisms contributing to positive student−teacher congruence effects in schools. With greater clarity, targeted policies and practices can be designed to have the greatest positive impact on the educational experiences of students belonging to marginalized groups. Findings have wide-ranging implications for how online course vendors design and school districts select online courses, as well as what policy interventions should be implemented to improve the academic experiences and outcomes of students belonging to marginalized groups.

Method

Online Instructional Environment

The courses studied were developed by one of the largest online course vendors in the United States. In communications with the vendor, they emphasized that all courses were developed by teams of experienced instructional designers in alignment with best practices including the iNACOL National Standards for Quality Online Courses. When asked

TABLE 1
Descriptive Statistics of Students in Each Analytic Sample

	Enrolled in top 30 course(s)	Course with gender diversity	Course with Black instructor	Course with Hispanic instructor	Cases with pretest
Female	0.448 (0.497)	0.451 (0.498)	0.435 (0.496)	0.458 (0.498)	0.448 (0.497)
Black	0.689 (0.463)	0.688 (0.463)	0.692 (0.462)	0.679 (0.467)	0.720 (0.449)
Hispanic	0.210 (0.407)	0.213 (0.410)	0.215 (0.411)	0.221 (0.415)	0.190 (0.392)
White	0.068 (0.252)	0.067 (0.250)	0.065 (0.246)	0.065 (0.247)	0.064 (0.244)
English language learner	0.121 (0.326)	0.122 (0.328)	0.125 (0.331)	0.128 (0.334)	0.103 (0.303)
Free/reduced price lunch eligible	0.812 (0.391)	0.815 (0.389)	0.824 (0.381)	0.825 (0.380)	0.824 (0.381)
Special education eligible	0.251 (0.433)	0.256 (0.436)	0.285 (0.452)	0.245 (0.430)	0.241 (0.428)
Prior year GPA	1.429 (0.795)	1.435 (0.780)	1.341 (0.771)	1.440 (0.810)	1.398 (0.770)
Failed course(s) in prior year	0.867 (0.339)	0.875 (0.331)	0.890 (0.313)	0.851 (0.356)	0.865 (0.331)
Pretested for one or more lesson	0.497 (0.500)	0.484 (0.500)	0.462 (0.499)	0.611 (0.488)	1.000 (0.000)
Number of students	7,328	6,129	4,278	2,663	3,575
Number of student–lesson observations	300,809	205,520	109,670	67,417	67,637

about how much discretion instructors were allowed, a spokesperson responded,

> [The vendor's] content development team provides clear guidelines, scripts, and talking points for online instructors. Because [the vendor]'s online instructors are highly qualified and state certified, the company respects instructors' experience and knowledge to make edits that ensure effective and natural lesson delivery. A team of editors and content experts review all online material before it is released. (Personal communication, July 8, 2019)

Each vendor-hired and trained remote instructors was provided scripts and talking points in addition to lecture slides that dictated lesson content and structure, and the research team observed minimal evidence of variation in conversational styles or ad-lib that might indicate that instructors were enacting discretion in how they taught (Darling-Aduana & Hemingway, 2021). The primary responsibilities of the remote instructor were to deliver a lecture according to the vendor-developed script during which they addressed students directly (but asynchronously) and otherwise presented themselves as the students' teacher and an expert in the subject matter covered. The form, content, and delivery of the lecture were designed to mirror a traditional, teacher-directed classroom. There was no means to facilitate communication between the remote instructor and the students viewing the prerecorded lectures.

Most students within the district were assigned to online course taking for credit recovery. Students accessed the asynchronous, online course within a blended hybrid model where they were assigned a class period and computer lab class location to make progress on course content (Christensen et al., 2013). In addition to their assigned lab period, students could make course progress anytime, anywhere. Students with limited access to devices or Internet access at home could make progress on their phones or at the local library, which partnered with the school district to provide access to computers for students enrolled in online courses.

The district staffed the computer labs used for online course taking with lab monitors who spent most of their time resolving technical issues and administrative issues. However, even these types of student–lab monitor interactions were infrequent. In observations of the physical computer lab setting, I identified only two minutes of interaction between the lab instructor and students during a typical 40-minute class period. The primary instructional responsibilities of these lab monitors were grading the occasional open-ended student response using a vendor-provided rubric; most items were graded automatically by the online course system.

Data and Sample

This study was conducted as part of a larger Research-Practice Partnership (Heinrich et al., 2019) and relies on administrative data provided by a large, urban school district in the Midwest that serves a predominately low-income, minoritized student population. As shown in Table 1, 81% of students in the analytic sample qualified for FRL, 69% identified as Black, and 21% identified as Hispanic, which largely reflects the demographics of the district. Data were provided for the 2016–2017 and 2017–2018 school years for all 9th through 12th-grade students. The district provided data on student achievement and sociodemographic variables. Among students enrolled in online courses, I also had access to information on course-taking behaviors, including which (and how many) online courses they enrolled in, how many times a student logged into the course system, and any assessment scores associated with each login. Over the 2-year period for which data were provided, 300,809 student–lesson observations (representing 7,328 unique students) were available for analysis with anywhere from 20 to 60 lessons per course.

Due to the time-intensive nature of coding the racial/ethnic and gender identity of the remote instructor for each

TABLE 2
Gender and Racial/Ethnic Diversity by Course

Course	Gender diversity	Racial/ethnic diversity	Description
Algebra 1 A	Yes	Yes	1 Black instructor
Algebra 1 B	Yes	—	
Algebra II A	Yes	Yes	1 Hispanic instructor
Algebra II B	Yes	—	
Biology A	Yes	Yes	1 Black instructor
Biology B	Yes	—	
Career Development	—	—	
Chemistry A	Yes	—	
Chemistry B	Yes	—	
Citizenship A	—	—	
Citizenship B	Yes	—	
ELA 9 A	Yes	Yes	1 Black instructor
ELA 9 B	Yes	Yes	2 Black instructors, 1 Hispanic instructor
ELA 10 A	—	—	
ELA 10 B	—	—	
ELA 11 A	Yes	—	
ELA 11 B	Yes	Yes	1 Black instructor
ELA 12 A	—	—	
ELA 12 B	—	—	
Geometry A	Yes	Yes	1 Hispanic instructor
Geometry B	Yes	—	
Healthy Living	—	Yes	1 Hispanic instructor
Entrepreneurship	—	—	
Personal Finance	Yes	Yes	1 Black instructor, 1 Hispanic instructor
Physical Science A	Yes	Yes	1 Black instructor
Physical Science B	Yes	—	
U.S. History A	Yes	Yes	2 Black instructors
U.S. History B	Yes	Yes	3 Black instructors
Word History A	Yes	Yes	3 Black instructors
World History B	Yes	Yes	1 Black instructor

Note. ELA = English and language arts. For year-long courses, "A" signifies the first semester and "B" signifies the second semester of a course.

lesson, I limited the analytic sample to the 30 courses in which the most students enrolled over the study period, which represented 80% of all online courses enrolled in during the study period. Each course contained lessons taught by anywhere from 1 to 17 unique remote instructors. Sometimes a single instructor taught an entire unit (containing 4–5 lessons) in a row. Other times, the instructor changed from lesson to lesson. In total, I identified 161 remote instructors, of which 139 presented as White, 17 as Black, and 5 as Hispanic.

Of the 30 online courses, I identified variation in the race/ethnicity of the instructors in 14 courses and gender variation in 22 courses. Coding of the race/ethnicity and gender of remote instructors incorporated the assessment of two raters based on characteristics observable in the lecture videos. Disagreements between raters were resolved through discussion and the assistance of a third rater. I refer to the raters' evaluations as presenting identity because it reflects how

students would likely categorize the instructor's identity but may not in all cases reflect how the instructor would have self-identified. While there is merit to self-identification, presenting identity is likely more relevant to student perceptions of an instructor's identity as examined in this study. Additional information on racial/ethnic and gender diversity by course is available in Table 2.

Thus, while I present descriptive statistics on data from all 30 courses, analyses requiring teacher racial/ethnic or gender diversity within the same course were limited to the courses that met the diversity precondition. Models examining student–instructor racial congruence among students identified as Black were limited to 11 courses. Further limiting this analytic sample to only cases of students identified as Black resulted in 68,880 student-by-lesson cases of 2,801 students. Analysis examining student–instructor racial congruence among students identified as Hispanic was limited

to five courses. Further limiting this analytic sample to only cases of students identified as Hispanic resulted in 14,029 student-by-lesson cases of 531 students. Among the 22 courses with variation in teacher gender, there were 205,520 student–lesson observations representing 6,129 unique students, with a little over half those cases representing students identified as male.

Outcome Measures

Outcome variables of interest included the first score the student earned on the end-of-lesson quiz, whether the student attempted the entire lesson (vs. starting and not finishing), and the time students logged in each fully attempted lesson. End-of-lesson scores were based on student responses to a 10-item, predominately (or entirely) multiple-choice quiz. Most questions required remembering and reciting facts or processes covered in the lecture video. The average score on the end-of-lesson quiz was 85% ($SD = 14$) with an observed range of 0% to 100%. Average quiz scores by lesson varied from 70% to 100% with all students receiving a 100% on their first quiz attempt on 12 of the 966 lessons. Three, five, and eight of these instances were in courses with instructor diversity by race, ethnicity, or gender, respectively. To adjust for the left skew in the data and varied standard deviations across lessons, I standardized the quiz scores within each lesson before estimating models. As a result, this transformed end-of-lesson quiz score had a mean of zero and a standard deviation of one.

The Fully Attempted Lesson measure is an indicator variable for whether a student completed all tasks associated with a lesson. Students did not need to earn a passing grade to receive credit for attempting the full lesson, but most did eventually. As such, whether students attempted the full lesson serves as a measure of persistence not conditional on achievement. In addition, examining this measure allows for the examination of student-by-lesson observations where the student accessed the lesson but did not complete the end-of-lesson quiz, resulting in a larger sample. Since the prerecorded video was the first task, students were expected to complete in a lesson[1], any subsequent persistent or disengagement occurred after students knew the instructors' presenting identities. In total, students attempted all elements of the lesson in 86% of cases ($SD = 35$).

Active time provided a measure of behavioral engagement for students that was based on the number of minutes the student spent logged into the online course system that the system recorded as making progress toward completing the lesson. For instance, time spent reviewing previously completed activities or the course map was excluded from this measure. I limited the examination of active time to cases where students attempted all elements of the lessons, including the end-of-lesson quiz, to prevent incomplete sessions from biasing estimates. I also purposefully excluded

time devoted to watching the video lecture, which would introduce noise related to the length of the lecture video from the measure of active time. Active time per lesson ranged from 0 to 101 hours with a median of 2.3 and a mean of 3.6 hours per lesson. The large right tail is indicative of a problem common with this type of metric, in that the system might incorrectly identify students as completing course content when they are away from or otherwise not engaging with the course. In general, there was a small negative association between active time and first quiz score, with students completing the lesson in less time being more likely to score well on their first quiz attempt and thus more likely to avoid having to retake the lesson. However, among students who logged fewer than one hour per session, increased active time was associated with higher quiz scores.

To address some of the inconsistencies and limitations of the active time measure, I transformed the variable prior to using it as a dependent variable. In recognition of variation in the designed length of each lesson, I centered the measure of active time within each lesson. Furthermore, due to the right tail in the active time variable, I transformed and used the natural log of the centered measured of active time as my dependent variable. This changed the interpretation of coefficients from a point increase to a percent increase. As a robustness test, I also estimated models dropping the top 1% and 5% of cases. Results remained qualitatively similar to the main estimates when data were trimmed in this manner.

Analytic Strategy

Using a student-by-course fixed effect strategy, I examined associations between lesson outcomes and whether the student experienced racial/ethnicity or gender congruence with the remote instructor for a given lesson. This strategy allowed for the exclusion of all variance associated with student and course-specific information that was constant during the semester that the student completed the course. Thus, the use of student-by-course fixed effects provided a more plausibly causal estimate than using covariate-adjusted ordinary least squares due to the removal of unobserved student and course characteristics that were fixed over time and might be related to both course assignment and outcomes. However, as stated previously, the use of this strategy focused exclusively on courses where there was variation in the type of instructor identity being examined.

The estimation strategy is summarized in the equation below. I estimated the model separately for each racial/ethnic and gender group, limiting the analytic sample to students who met the condition of interest enrolled in courses with variation in teachers based on that same condition.

$$\gamma_{icl} = \beta_0 + \beta_1 inst\_identity_{icl} + \boldsymbol{A_{cl}}\beta + \boldsymbol{S_{icl}}\beta + \alpha_{ic} + \varepsilon_{icl} \quad (1)$$

The use of student-by-course fixed effects (α_{ic}) allowed for the examination of associations with student–instructor racial/ethnicity or gender congruence ($inst\_identity_{icl}$) for a given student (i) from lesson to lesson (l) within a given course (c). The equation controls for a vector of lesson (A) covariates, including the lesson order in the course and binary indicators for whether various activities were present in the lesson. These activities included a quiz (observed in 95% of lessons), warm-up exercise (56%), assignment (53%), review component (51%), keyword definitions (29%), outside electronic resource (4%), or accompanying text (4%). These activities served as a proxy for lesson expectations and difficulty that are likely to be related to both end-of-lesson quiz scores and the amount of time required to complete lesson content. I controlled for these activities, as curriculum experts—and not the remote instructors—choose how to design the course and which activities to include. I also controlled for a vector of student covariates (S) including whether the student took the pretest for a given lesson, which provided the opportunity to test out of some lesson content, as well as the score received on the pretest where applicable. All models used robust standard errors, clustered at the student-by-course level, to account for correlations between the error terms.

In addition, I ran several sensitivity tests to demonstrate robustness to alternative model specifications. These tests include estimating models with a teacher fixed effect as well as conditioning on remote instructor frequency and the authenticity of lesson tasks. The teacher fixed effect model removes variation associated with each teacher that is constant across all lessons that they teach, thus more effectively controlling for factors such as differential instructor quality or systemic assignment to certain lessons by identity. Despite these benefits, I chose not to make these models my preferred specifications due to the limited number of courses where more than one teacher of a minoritized racial/ethnic group taught. Nonetheless, consistent results between models would minimize concerns regarding bias in the main estimates due to differential teacher quality or systemic assignment to courses.

The model controlling for teacher frequency tests the hypothesis that students might be responding differentially to teachers based not on their identity but on increased familiarity with teachers who teach more lessons in the course. Last, I measured authenticity using the Authentic Online Work Rubric, an observation tool that allows for the generation of two standardized, continuous variables measuring the extent to which each lesson facilitated higher order thinking and provided real-world relevance (Darling-Aduana, 2021). Since lesson tasks were dictated by curriculum developers versus the remote instructors, controlling for lesson authenticity provided additional evidence regarding whether it was more likely that students were responding to variation in teacher identity versus variation in lesson characteristics associated with teacher identity due to systemic teacher assignment. Additional information on the motivation for each of these tests, as well as specifics on how they were conducted, is available in the appendix.

In addition to the average treatment estimates specified above, I examined heterogeneous effects using Equation 1 by limiting the sample to each subject or subgroup of interest. Based on my review of prior research, I examined estimates for each of the four core subjects (ELA, social studies, math, and science; Keiser et al., 2002; Solanki & Xu, 2018; Xu & Li, 2018), as well as for students who qualified for FRL and students whose prior year GPA was below a 2.0 (indicating below a C average; Dee, 2004, 2005; Egalite et al., 2015; Joshi et al., 2018). To examine variation by student intersectionality, I explored racial/ethnic congruence limiting the sample to first female and then male students. Last, it is important to note that potential sources of bias may remain in the estimates presented below. As such findings should be interpreted as descriptive rather than causal.

Summary of Findings

Student and Instructor Demographic Identities

As shown in Figure 1, 7% students in the analytic sample identified as White and 45% as female compared with 87% of remote instructors who presented as White and 57% that presented as female.[2] The in-person lab instructors hired by the school district more closely mirrored the student population in terms of racial and ethnic diversity, although discrepancies remained. Consequently, students identified as Black and Hispanic were each taught by a remote instructor of the same race/ethnicity in only 5% of student–lesson observations (representing approximately 10,000 and 3,000 cases, respectively). In contrast, students identified as White were taught by an instructor of the same race in 87% of observations.

The proportion of instructors who identified as Black or Hispanic varied slightly by subject, as shown in Figure 2. I identified the largest percentage in elective, ELA, and social studies courses (28%, 12%, and 10% lessons, respectively) compared with only 5% of instructors in math and 7% in science. The gender distribution of instructors also varied by subject. In ELA, 83% of instructors presented as female, with 43% to 62% of instructors presenting as female in other subject areas.

Student–Remote Instructor Identity Congruence

Next, I examined lesson outcomes by remote instructor characteristics using a student-by-course fixed effect strategy. Estimates from these models are presented in Table 3. As a reminder, sample sizes change between models examining different identities, because I restricted models to cases from students belonging to the identity group of
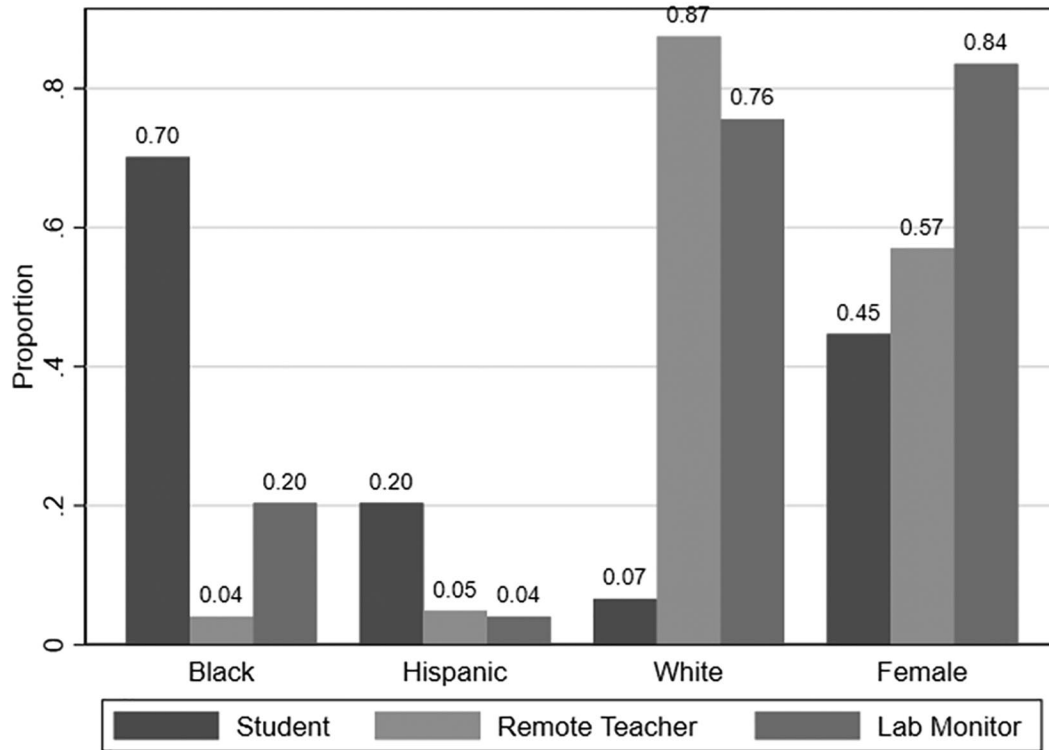
FIGURE 1. *Student and teacher racial/ethnic and gender proportions.*

interest and to courses where there was variation in the teacher identity examined. I observed no significant associations between quiz scores or whether the student attempted the entire lesson and congruence with the remote instructor in the base model (column 1). In addition to the lack of significance, the magnitude of estimates was small for students belonging to minoritized groups (less than or equal to 0.01 standard deviations on the end-of-lesson quiz and 0.4% for attempting the entire lesson). The small magnitude and lack of significant associations held across various model specifications, including the introduction of a teacher fixed effect (column 2), controlling for instructor frequency (column 3), and conditioning on the authenticity level of lesson tasks (column 4).

In contrast, I identified a significant positive association between student–instructor racial/ethnic match and active time for students identified as Black and Hispanic. These students logged 1.5% to 2.0% more active time when taught by a same-race/ethnicity remote instructor. To place these findings in context, I conducted two exploratory supplemental analyses, a quantile regression of the main model by prior year GPA (based on the assumption that students' default engagement patterns might vary across this dimension) and a series of linear probability models that estimated the likelihood that students would switch from one active

time category to another. The estimates from these models suggested that the increase in active time identified among students identified as Black was most often an increase from minimal active time to a small to moderate amount of active time (totaling less than 42 minutes, the typical length of a class period in the district). This type of increase was associated with higher end-of-lesson quiz scores. Furthermore, the improvements in active time were most pronounced among students with prior year GPAs in the bottom quartile (0.0–0.8 prior year GPA), who were also the quartile most likely to log minimal active time for a given lesson. For students identified as Hispanic, a less positive pattern emerged, with increases in active time driven by increases in active time beyond 42 minutes. This type of increase was associated with lower end-of-lesson quiz scores.

Sensitivity Tests

Estimates remained comparable in magnitude (if not significance) when introducing a teacher fixed effect (Table 3, column 2), controlling for instructor frequency (Table 3, column 3), and conditioning on the authenticity of lesson tasks (Table 3, column 4). However, many of the shifts present when conditioning on authenticity—including the decreased
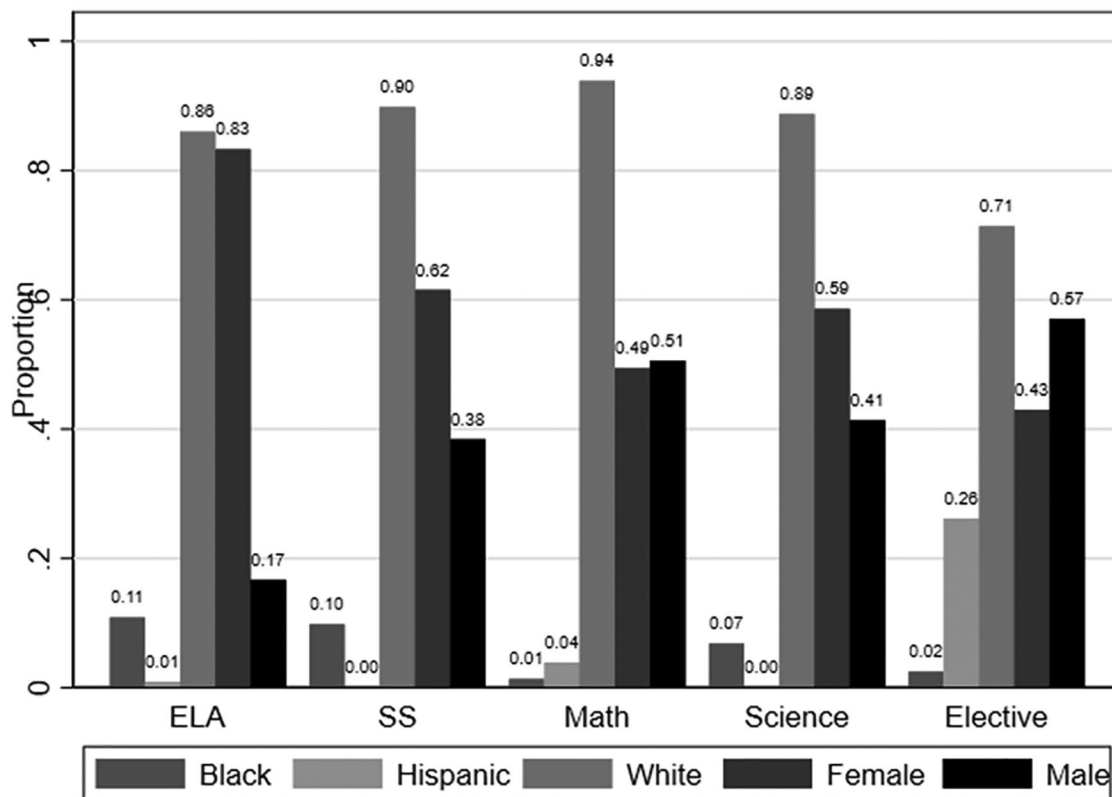
FIGURE 2.   *Remote instructor racial/ethnic and gender proportions by subject.*

magnitude of the association between active time and student–instructor congruence—were present when limiting the sample to the 10 courses with authenticity scores even without controlling for authenticity. As such, this deviation is more likely in response to the reduced sample versus variability due to differences in authenticity from lesson to lesson.

Variation by Subject and Subgroup

Prior literature in face-to-face instructional settings indicated that students were likely to respond differently to racial/ethnic or gender congruence with their instructor depending on the course subject and student characteristics (Dee, 2004, 2005; Egalite et al., 2015; Joshi et al., 2018; Keiser et al., 2002; Solanki & Xu, 2018; Xu & Li, 2018). I examined whether there was evidence of these heterogeneous effects in online courses in Table 4 for all subgroups of interest except for students identified as Hispanic, which I was unable to examine due to low instances of student–instructor congruence when disaggregating the sample. Consistent with estimates from the main model specifications, few of the subgroup analyses identified significant associations between student–instructor congruence and

either end-of-lesson quiz score or fully attempting a lesson apart from a few subject-specific patterns. In science, students presenting as Black were 1% more likely to fully attempt a lesson when taught by a same-identity remote instructor. However, there were only two remote science instructors identified as Black who taught a total of 11 classes, so this finding should not be overinterpreted. Additionally, students identified as female were 1.6% more likely to fully attempt a science lesson when taught by a same-identity instructor, and students identified as male scored 0.03 standard deviations higher in social studies when taught by a same-identity instructor. The positive association observed among students identified as male in social studies is particularly interesting, because prior research has focused on positive benefits for women in STEM fields but less so on benefits for men in fields historically considered more feminine. I discuss the implications of this finding in greater detail below.

Additionally, evidence of a possible association between active time and student–instructor congruence among students identified as Black persisted across many subjects (ELA and social studies) and subgroups (FRL, < 2.0 prior year GPA, female, and male students). In contrast, students

TABLE 3
Lesson Achievement, Persistence, and Engagement by Student–Instructor Match

	(1)	(2)	(3)	(4)
	Base model	Teacher fixed effect	Instructor frequency	Authentic work
Panel A: End-of-lesson quiz score (Std.)				
Black match	−0.011 (0.012)	−0.009 (0.013)	−0.008 (0.012)	−0.018 (0.017)
N	62,075	62,075	62,075	25,554
Hispanic match	0.009 (0.027)	—	0.012 (0.028)	0.002 (0.046)
N	12,199		12,199	4,800
White match	−0.028 (0.027)	−0.012 (0.027)	−0.026 (0.028)	0.010 (0.039)
N	9,113	9,113	9,113	3,820
Male match	0.006 (0.008)	0.013 (0.008)	0.014* (0.008)	−0.001 (0.012)
N	89,814	89,814	89,814	38,898
Female match	−0.005 (0.008)	−0.004 (0.009)	−0.011 (0.009)	−0.008 (0.012)
N	78,411	78,411	78,411	34,086
Panel B: Fully attempted lesson				
Black match	0.004 (0.003)	0.005 (0.003)	0.005 (0.003)	0.001 (0.004)
N	68,880	68,880	68,880	28,564
Hispanic match	0.000 (0.005)	—	−0.000 (0.005)	0.014 (0.011)
N	13,907		13,907	5,722
White match	0.003 (0.006)	0.006 (0.006)	0.003 (0.006)	−0.002 (0.009)
N	10,028	10,028	10,028	4,289
Male match	−0.001 (0.002)	−0.002 (0.002)	0.000 (0.002)	0.005* (0.003)
N	100,645	100,645	100,645	43,917
Female match	0.003 (0.002)	0.004* (0.002)	0.002 (0.002)	−0.001 (0.003)
N	88,388	88,388	88,388	38,739
Panel C: Log of active time (centered)				
Black match	0.015*** (0.005)	0.014** (0.006)	0.017*** (0.005)	0.007 (0.007)
N	62,075	62,075	62,075	25,554
Hispanic match	0.020*** (0.006)	—	0.013 (0.011)	0.028*** (0.008)
N	12,199		12,199	4,800
White match	−0.006 (0.006)	−0.005 (0.007)	−0.007 (0.006)	−0.000 (0.009)
N	9,113	9,113	9,113	3,820
Male match	0.003 (0.003)	0.005 (0.003)	0.004 (0.004)	−0.001 (0.004)
N	89,814	89,814	89,814	38,898
Female match	−0.002 (0.003)	−0.006 (0.004)	−0.003 (0.004)	0.006 (0.004)
N	78,411	78,411	78,411	34,086

Note. I was not able to estimate a model with instructor fixed effects for Hispanic instructors, as there were no courses with more than one Hispanic instructor. Each cell represents estimates from a different model.
*.10, **.05, ***.01 significance level.

identified as White logged 5.4% less active time in math and 1.1% less time if they qualified for FRL when taught by a same-race instructor. At the same time, students identified as male with prior year GPAs below 2.0 logged 0.9% more active time when taught by a male instructor.

Contributions and Limitations

This is one of the first studies to examine the characteristics of online course instructors. Beyond establishing demographics, the unique nature of online course taking allows for greater separation of any possible educational benefits of teacher versus student-based mechanisms related to representation than feasible within a traditional, face-to-face instructional setting. Another strength of this study is the relatively large sample size and detailed data that allow for the use of a student-by-course fixed effect strategy. Furthermore, while concerns regarding generalizability may remain to other online contexts, student populations, and online courses in which no racial/ethnic variation in instructor identity was observed, claims of generalizability are strengthened by the examination of content developed by one of the largest online course vendors in the United States.

TABLE 4
Student−Teacher Congruence by Subject and Subgroup

	ELA	SS	Math	Science	FRL	<2.0 Prior year GPA	Female	Male
Panel A: End-of-lesson quiz score (std.)								
Black match	−0.028 (0.021)	−0.002 (0.020)	−0.049 (0.053)	0.016 (0.027)	−0.016 (0.013)	−0.018 (0.014)	−0.014 (0.017)	−0.008 (0.018)
N	13,338	27,732	6,499	8,747	50,657	43,282	30,447	31,628
White match	−0.013 (0.047)	−0.029 (0.068)	0.064 (0.109)	0.073 (0.087)	0.008 (0.033)	−0.023 (0.038)	−0.006 (0.050)	−0.042 (0.031)
N	1,498	2,292	2,348	753	5,913	5,364	3,599	5,514
Male match	0.026 (0.017)	0.026** (0.013)	−0.024 (0.016)	0.003 (0.024)	−0.001 (0.009)	0.006 (0.009)	—	—
N	15,600	25,386	30,282	13,980	67,206	66,769	—	—
Female match	0.002 (0.018)	0.004 (0.015)	−0.005 (0.017)	−0.030 (0.027)	−0.006 (0.010)	−0.004 (0.010)	—	—
N	11,125	22,844	28,048	12,078	62,576	48,461	—	—
Panel B: Fully attempted lesson								
Black match	0.006 (0.005)	0.004 (0.005)	−0.029 (0.018)	0.013** (0.006)	0.004 (0.003)	0.005 (0.003)	0.002 (0.004)	0.005 (0.004)
N	14,660	29,560	7,341	10,579	56,364	48,025	33,733	35,147
White match	0.003 (0.014)	−0.011 (0.014)	0.020 (0.021)	0.010 (0.012)	0.003 (0.007)	0.006 (0.009)	0.010 (0.009)	−0.002 (0.007)
N	1,656	2,407	2,672	879	6,544	5,957	3,987	6,041
Male match	0.000 (0.004)	0.002 (0.003)	−0.001 (0.004)	−0.005 (0.005)	−0.001 (0.002)	−0.001 (0.002)	—	—
N	17,017	26,979	34,449	16,881	75,697	74,777	—	—
Female match	0.004 (0.006)	−0.001 (0.004)	0.003 (0.004)	0.016*** (0.005)	0.001 (0.002)	0.003 (0.003)	—	—
N	12,159	24,395	32,169	14,597	70,747	54,521	—	—
Panel C: Log of active time (centered)								
Black match	0.016** (0.008)	0.023*** (0.009)	−0.001 (0.019)	−0.006 (0.013)	0.017*** (0.006)	0.014** (0.006)	0.013* (0.007)	0.010* (0.006)
N	13,338	27,732	6,499	8,747	50,657	43,282	30,447	31,628
White match	−0.010 (0.012)	−0.016 (0.010)	−0.054*** (0.016)	0.005 (0.023)	−0.011** (0.006)	−0.004 (0.007)	−0.002 (0.007)	−0.010 (0.006)
N	1,498	2,292	2,348	753	5,913	5,364	3,599	5,514
Male match	0.008 (0.005)	0.003 (0.007)	−0.003 (0.006)	0.009 (0.007)	0.002 (0.004)	0.009*** (0.004)	—	—
N	15,600	25,386	30,282	13,980	67,206	66,769	—	—
Female match	0.004 (0.006)	0.007 (0.006)	−0.008 (0.007)	−0.012 (0.009)	−0.000 (0.004)	−0.000 (0.004)	—	—
N	11,125	22,844	28,048	12,078	62,576	48,461	—	—

Note. ELA = English and language arts; SS = social studies; FRL = Free or Reduced-Price Lunch. I was not able to examine differential effects for students identified as Hispanic due to low instances of student−teacher congruence when disaggregating the sample. Each cell represents estimates from a different model.
*.10, **.05, ***.01 significance level.

Despite these strengths, there are also limitations to the study design that should be considered when interpreting results. Perhaps, most important, estimates were not causal, as there might remain associations between instructor race/ethnicity or gender and instructional content relevant to student outcomes that I was not able to identify and control for due to systematic assignment of instructors to teach lessons by identity. It is also possible, especially considering the low proportion of minority instructors, that there might be something systematically more difficult about getting hired to be an online course instructor for minority instructors. This might result in higher qualified minority instructors being hired, on average, than their White counterparts that faced fewer employment barriers. Furthermore, the low proportions of instructors presenting as a racial/ethnic minority reduced the effective sample size.

This analysis also rests on the assumption that a single lesson is sufficient time for any benefits associated with representation to occur. While, racial recognition and the triggering of stereotype threat are instantaneous, often automatic processes, the identification of an instructor as a role model may require long-term opportunities to interact, and thus may be less likely to be detected in the learning environment studied. Conversely, role model effect can also be triggered by an individual whom a student has never met (i.e., Barack Obama; Morgenroth et al., 2015), which could be more easily fostered in the asynchronous, prerecorded structure studied. There also remains the possibility that it is not the number or length of exposures but the virtual classroom environment in and of itself that fails to produce a student response. Specifically, students might not benefit (or benefit less) from representation in an asynchronous, online instruction setting because stereotype threat or role model effects might be less likely to be triggered when because students know that the prerecorded instructor cannot respond differentially to them based on their identity.

Furthermore, I was limited in my selection of dependent variables to the use of lesson-level achievement and engagement measures captured by the online course system. The end-of-lesson quiz was relatively short and tended toward easier questions, resulting in potentially nosier estimates than when using a psychometrically validated standardized assessment. The precision of resulting estimates minimizes some of these concerns but does not preclude the possibility that any effect of student–instructor congruence may have been identified if more complex tasks were assessed.

There is a similar concern with the measure of active time collected by the online course system. The large right-hand tail indicates that the system was likely unable to parse out some types of disengagement in its measure of active time. This introduces noise and complicates the interpretation of resulting estimates. However, the robustness of findings to models trimming the top 1% and 5% of active time, as well as models excluding students recorded as logging no active time, minimizes concerns that the primary models were biased by these outliers. In fact, supplemental analyses provided suggestive evidence that increased active time among students identified as Hispanic might have been biased by imperfect measurement of active time within the online course system that sometimes resulted in logging a likely inflated number of active hours. Conversely, the same analyses suggested that increased active time among students identified as Black when taught by a same-race instructor was likely driven by active time changes in the more moderate active time range, which were more likely to represent true increases in engagement.

Additionally, it is also not possible to directly compare effect sizes from this study with those identified in prior studies, as most prior studies used general content area standardized assessments instead of the more closely aligned content-specific end-of-lesson quizzes examined in this study. Relatedly, in light of the relatively modest magnitude year-long effects between student–instructor congruence and achievement, if lesson-level associations were proportionally smaller in magnitude, it might be infeasible to estimate models with sufficient precision to identity potentially very small associations.

Last, the sample for this study included students enrolled in online courses predominately for the purpose of credit recovery within a large, urban, low-resourced school district. The experiences of these students are of great importance to researchers and practitioners interested in educational equity. At the same time, this constraint should be recognized when interpreting and attempting to generalize any findings to a wider population of students. It is also important to note that prior research finds that representation effects are often largest among students belonging to marginalized groups due, potentially, to fewer alternative pathways to relevant social and cultural capital (i.e., Grissom et al., 2015). As such, any identified associations among the student population studied are likely larger in magnitude than if a less marginalized sample were studied.

Discussion

Within the large, urban school district in which the study occurred, I identified large differences in the racial/ethnic and gender distribution of the remote instructors compared to the student population served that far exceeded existing disparities among the lab monitors hired by the school district studied. When examining the extent to which being taught by a remote instructor of the same race/ethnicity or gender was associated with course outcomes, I identified few significant associations with end-of-lesson quiz scores or whether students fully attempted the lesson. However, students identified as Black or Hispanic logged more active time in lessons taught by a same-race/ethnicity instructor. Supplemental analyses suggested that the increase among

students identified as Black, in particular, was likely to be associated with increased engagement with lesson tasks.

These estimates were robust to several sensitivity tests, including those accounting for teacher fixed effect and remote instructor frequency. Additionally, estimates were consistent in directionality and magnitude (if not significance) when accounting for the authenticity of lesson tasks, which was only available for a subset of courses. Notable findings from the examination of associations by subject and subgroup included the prevalence across marginalized subgroups of the positive association between race match and active time among students identified as Black. This identification of more pronounced positive outcomes for students identified as Black when experiencing student–instructor congruence is consistent with findings identified in prior research and is likely due to the greater systemic inequities faced by students within this group (i.e., Gershenson et al., 2017; Grissom et al., 2015; Joshi et al., 2018). In addition, by gender, students identified as male scored slightly higher in social studies when taught by a same-gender teacher, while students identified as female (and Black) were slightly more likely to fully attempt science lessons when taught by a same-gender (or same-race) teacher, respectively.

Generally, null results when examining end-of-lesson quiz scores and whether students fully attempted lessons after watching the initial video lecture could indicate that student-based mechanisms associated with representation were not triggered in the type of impersonal setting studied, where students and teachers did not interact. However, the identification of a consistent, positive association between being taught by a same-race instructor and active time among students identified as Black appears to contradict this possibility. Furthermore, prior research has identified differential responsiveness in online instructional settings based on something as small as a name with certain racial/ethnic and gender connotations influenced behavior (Baker et al., 2018). Instead, the lack of an identified relationship (if one does exist) might have more to do with the measurement error associated with the short, relatively easy end-of-lesson quizzes used to measure achievement. Similarly, fully attempting a lesson was a relatively low benchmark for engagement.

As such, driven by associations with active time, findings support (at least in part) the hypothesis that passive forms of representation account for some variation in the engagement of students belonging to minoritized groups even in the fully standardized, asynchronous online course environment studied. These associations persist—and often increase in magnitude—when examining subgroups experiencing additional forms of marginalization (i.e., qualifying for FRL). Furthermore, I also identified some positive associations between course outcomes and being taught by a same-gender instructor in subjects more often associated with opposite-gender attributes. Therefore, based on the potentially

diluted and more limited pathways for even passive representation to occur in the online environment studied, it is likely that passive mechanisms also drove at least part of the benefits of student–teacher congruence identified in prior literature conducted in face-to-face settings (i.e., Dee, 2004, 2005; Egalite et al., 2015; Gershenson et al., 2017; Solanki & Xu, 2018). In particular, this speaks to the likely role of passive mechanisms associated with shorter term, more automatic same-identity recognition that are not conditional on interaction.

Implications for Research, Policy, and Practice

This study contributes to the larger literature on representation within education for students belonging to marginalized groups (i.e., Egalite et al., 2015; Gershenson et al., 2017; Grissom et al., 2015). Compared with staffing traditional classrooms, it is relatively easier to provide students access to instructors belonging to minorized racial/ethnic groups through online courses due to the standardized nature of the lesson. Instead of positively influencing a single class of 30 students, increasing the diversity of instructors in even a single course developed by the online course vendor studied has the potential to affect the academic experiences of tens of thousands of students across the country. Despite this potential, I identified less racial/ethnic diversity among remote instructors within the most frequently accessed courses than among in-person lab monitors hired by the district studied. Beyond speaking to the likely implications of diversifying remote instructors, these findings also provide insight on expected improvements (or lack of improvements) in student outcomes educators can expect by solely focusing on diversifying representation without also diversifying content, frameworks, and perspective in instructional materials, such as textbooks, more broadly.

Specifically, the use of online session-level course data allowed for the identification of within student and course variation by student–instructor congruence. Findings demonstrated that positive associations identified in prior literature conducted in traditional, face-to-face settings among students identified as Black might be due at least in part to passive forms of representation, such as a reduction in stereotype threat or increase in role model effect, in addition to factors associated with active representation, such as through teacher discretion. Accordingly, continuing to expand representation in online courses appears merited due to the low relative cost even considering relatively small magnitude associations with engagement and generally null associations with achievement. This diversification is likely to have the most positive effect if combined with increased discretion for remote instructors belonging to a minoritized group (Fish, 2017; Grissom et al., 2009, 2015; Lindsay & Hart, 2017). There might also be benefits beyond the dependent variables examined in this study. For instance, normalizing

instructors of different identities as experts across subject areas could have broader societal benefits for all students (i.e., Paris, 2012). And in fact, I observed some negative associations between active time and student–instructor congruence among students identified as White in math and when also qualifying for FRL. This finding could reflect hiring practices that resulted in differential instructor quality by identity and/or it might reflect the presence of more salient identities for students identified as White in the sample, a majority of whom had struggled academically and/or qualified for FRL.

Researchers have called for a range of policy interventions to facilitate passive representation mechanisms including increasing the diversity of the teacher workforce (to facilitate role model effects and coproduction inducement) and minimizing teacher implicit bias (to reduce the triggering of stereotype bias) (Dee, 2005; Egalite et al., 2015; Gershenson et al., 2016, 2017; Grissom et al., 2015; Irizarry, 2015; Joshi et al., 2018; Lindsay & Hart, 2017). Beyond these recommendations, the identification of an association between student–instructor congruence and student engagement in the prerecorded, asynchronous online course environment studied also suggests that exposure to a same-identity expert virtually in low doses might be sufficient to experience some benefit. For instance, in addition to working to expand teacher workforce diversity, short-term solutions such as integrating guest speakers from diverse backgrounds into classroom instruction might also induce improved engagement.

Despite not being tested in this study, interventions listed above designed to enhance benefits associated with passive representation are likely to have an even bigger impact when combined with active mechanisms associated with teacher discretion, including integrating culturally relevant pedagogy and reevaluating how students are identified for gifted coursework, special education services, and disciplinary referrals (Baker et al., 2018; Burgess & Greaves, 2013; Carlana, 2019; Cherng & Halpin, 2016; Chin et al., 2020; Dee & Penner, 2019; Fish, 2017; Grissom et al., 2009, 2015; Lindsay & Hart, 2017). Future research should continue to disentangle the effects of mechanisms associated with teacher discretion to provide more nuanced and causal evidence regarding which policies are most likely to replicate the positive academic outcomes associated with exposure to teachers belonging to marginalized groups in both online and traditional, face-to-face instructional settings.

Appendix

Sensitivity Tests

To test the robustness of estimates from my preferred model, I conducted several sensitivity tests. For instance, I was concerned that the estimates from the primary model might have been biased if vendor hiring practices resulted in differential instructor quality by race/ethnicity or gender or students responded differently to the remote instructors for some other reason not related to congruence. For that reason, I estimated the following model, which included a teacher fixed effect (δ_{itc}) to prevent variance in teacher-specific characteristics from biasing estimates.

$$
\begin{aligned}
\gamma_{itcl} = \beta_0 + \beta_1 inst\_identity_{itcl} + A_{cl}\beta + S_{icl}\beta \\
+ \delta_{itc} + \alpha_{ic} + \varepsilon_{itcl}
\end{aligned}
\tag{2}
$$

The disadvantage of this model is that the main treatment estimates were based primarily on data from courses that were taught by two or more instructors of the identity of interest. As shown in Table 2, I identified only four courses taught by two or more instructors presenting as Black and none taught by two or more instructors presenting as Hispanic. Nonetheless, if estimates remained consistent after the inclusion of a teacher fixed effect, it would reduce the possibility that the main estimates were biased by endogenous teacher-level factors.

I was also concerned that any identified estimates might instead be attributed to greater familiarity with a particular remote instructor. To account for this increased familiarity, I estimated the following model that controlled for the number of lessons in a course each remote instructor taught ($\beta_2 inst\_freq_{tc}$).

$$
\begin{aligned}
\gamma_{itcl} = \beta_0 + \beta_1 inst\_identity_{itcl} + \beta_2 inst\_freq_{tc} \\
+ A_{cl}\beta + S_{icl}\beta + \alpha_{ic} + \varepsilon_{itcl}
\end{aligned}
\tag{3}
$$

If the inclusion of this covariate resulted in comparable estimates, it would indicate that the frequency of exposure to a particular remote instructor did not appear to be associated with changes in students' responsiveness to racial/ethnic or gender congruence.

Beyond controlling for teacher characteristics, I was concerned that there might remain endogenous lesson-level characteristics in terms of the content or type of tasks completed that might bias estimates. To examine the relevance of this concern, I substituted the variable for instructor frequency in Equation 3 with variables measuring the authenticity of each lesson. Authenticity data were only available for the 10 most enrolled in courses in the district, resulting in a reduced sample (see Darling-Aduana, 2021, for more information on how authenticity was conceptualized and measured). Within those 10 courses though, I was able to control for the extent to which each lesson facilitated higher order thinking, provided real-world relevance, and the interaction between those two terms. Each authenticity subscale was measured on a continuous, standardized scale. As authenticity was dictated by the curriculum developers and not the remote instructor, accounting for associated

characteristics provided an additional examination of whether any previously identified effects might be due to the systematic assignment of teachers of a given race/ethnicity or gender to different types of lessons.

ORCID iD

Jennifer Darling-Aduana https://orcid.org/0000-0002-7940-5662

Notes

1. A few lessons had warm-up activities before the video lecture. However, these activities were short and required limited interactivity, meaning that even in these instances the video lecture was the first major task students were expected to complete.

2. In instances, where the percentage of students or teachers did not add up to 100%, the additional percentages represent individuals identified as a race/ethnicity not listed, such as Asian or Pacific Islander.

References

Baker, R., Dee, T., Evans, B., & John, J. (2018). *Bias in online classes: Evidence from a field experiment* (CEPA Working Paper No. 18-03). Stanford Center for Education Policy Analysis.

Burgess, S., & Greaves, E. (2013). Test scores, subjective assessment, and stereotyping of ethnic minorities. *Journal of Labor Economics*, *31*(3), 535–576. https://doi.org/10.1086/669340

Carlana, M. (2019). Implicit stereotypes: Evidence from teachers' gender bias. *Quarterly Journal of Economics*, *134*(3), 1163–1224. https://doi.org/10.1093/qje/qjz008

Cherng, H. Y. S., & Halpin, P. F. (2016). The importance of minority teachers: Student perceptions of minority versus White teachers. *Educational Researcher*, *45*(7), 407–420. https://doi.org/10.3102/0013189X16671718

Chin, M. J., Quinn, D. M., Dhaliwal, T. K., & Lovison, V. S. (2020). Bias in the air: A nationwide exploration of teachers' implicit racial attitudes, aggregate bias, and student outcomes. *Educational Researcher*, *49*(8), 566–578. https://doi.org/10.3102/0013189X20937240

Christensen, C. M., Horn, M. B., & Staker, H. (2013). *Is K-12 blended learning disruptive? An introduction to the theory of hybrids*. Clayton Christensen Institute for Disruptive Innovation.

Darling-Aduana, J. (2021). Authenticity, engagement, and performance in online high school courses: Insights from micro-interactional data. *Computers & Education*, *167*(July), Article 104175. https://doi.org/j.compedu.2021.104175

Darling-Aduana, J., & Hemingway, K. (2021). *Representation is not enough: Teacher identity and discretion in an asynchronous, scripted online learning environment* [Paper presentation]. AEFP. (Virtual conference due to COVID-19.)

Dee, T. S. (2004). Teachers, race, and student achievement in a randomized experiment. *Review of Economics and Statistics*, *86*(1), 195–210. https://doi.org/10.1162/003465304323023750

Dee, T. S. (2005). A teacher like me: Does race/ethnicity or gender matter? *American Economic Review*, *95*(2), 158–165. https://doi.org/10.1257/000282805774670446

Dee, T. S., & Penner, E. (2019). *My brother's keeper? The impact of targeted educational supports* (CEPA Working Paper No.19-07). Stanford Center for Education Policy Analysis. http://cepa.stanford.edu/wp19-07

Egalite, A. J., Kisida, B., & Winters, M. A. (2015). Representation in the classroom: The effect of own-race teachers on student achievement. *Economics of Education Review*, *45*, 44–52. https://doi.org/10.1016/j.econedurev.2015.01.007

Fish, R. E. (2017). The racialized construction of exceptionality: Experimental evidence of race/ethnicity effects on teachers' interventions. *Social Science Research*, *62*, 317–334. https://doi.org/10.1016/j.ssresearch.2016.08.007

Gemin, B., Paper, L., Vashaw, L., & Watson, J. (2015). *Keeping pace with K-12 digital learning: An annual review of policy and practice*. Evergreen Education Group.

Gershenson, S., Hart, C. M. D., Lindsay, C. A., & Papageorge, N. W. (2017). *The long-run impacts of same-race teachers* (Discussion Papers, No. 10630). Institute of Labor Economics (IZA).

Gershenson, S., Holt, S. B., & Papageorge, N. W. (2016). Who believes in me? The effect of student–teacher demographic match on teacher expectations. *Economics of Education Review*, *52*, 209–224. https://doi.org/10.1016/j.econedurev.2016.03.002

Grissom, J. A., Kern, E. C., & Rodriguez, L. A. (2015). The "representative bureaucracy" in education: Educator workforce diversity, policy outputs, and outcomes for disadvantaged students. *Educational Researcher*, *44*(3), 185–192. https://doi.org/10.3102/0013189X15580102

Grissom, J. A., Nicholson-Crotty, J., & Nicholson-Crotty, S. (2009). Race, region, and representative bureaucracy. *Public Administration Review*, *69*(5), 911–919. https://doi.org/10.1111/j.1540-6210.2009.02040.x

Heinrich, C. J., Darling-Aduana, J., Good, A. G., & Cheng, H. (2019). A look inside online educational settings in high school: Promise and pitfalls for improving educational opportunities and outcomes. *American Educational Research Journal*, *56*(6), 2147–2188. https://doi.org/10.3102/0002831219838776

Herrington, J., Oliver, R., & Reeves, T. C. (2006). Authentic tasks online: A synergy among learner, task and technology. *Distance Education*, *27*(2), 233–248. https://doi.org/10.1080/01587910600789639

Irizarry, Y. (2015). Selling students short: Racial differences in teachers' evaluations of high, average, and low performing students. *Social Science Research*, *52*, 522–538. https://doi.org/10.1016/j.ssresearch.2015.04.002

Joshi, E., Doan, S., & Springer, M. G. (2018). Student-teacher race congruence: New evidence and insight From Tennessee. *AERA Open*, *4*(4), 1–25. https://doi.org/10.1177/2332858418817528

Kao, G. (2004). Social capital and its relevance to minority and immigrant populations. *Sociology of Education*, *77*(2), 172–175. https://doi.org/10.1177/003804070407700204

Keiser, L. R., Wilkins, V. M., Meier, K. J., & Holland, C. A. (2002). Lipstick and logarithms: Gender, institutional context, and representative bureaucracy. *American Political Science Review*, *96*(3), 553–564. https://doi.org/10.1017/S0003055402000321

Lim, H. H. (2006). Representative bureaucracy: Rethinking substantive effects and active representation. *Public Administration Review*, *66*(2), 193–204. https://doi.org/10.1111/j.1540-6210.2006.00572.x

Lindsay, C. A., & Hart, C. M. (2017). Exposure to same-race teachers and student disciplinary outcomes for black students in North Carolina. *Educational Evaluation and Policy Analysis*, *39*(3), 485–510. https://doi.org/10.3102/0162373717693109

McLoughlin, C., & Oliver, R. (2000). Designing learning environments for cultural inclusivity: A case study of indigenous online learning at tertiary level. *Australasian Journal of Educational Technology*, *16*(1). https://doi.org/10.14742/ajet.1822

Morgenroth, T., Ryan, M. K., & Peters, K. (2015). The motivational theory of role modeling: How role models influence role aspirants' goals. *Review of General Psychology*, *19*(4), 465–483. https://doi.org/10.1037/gpr0000059

Nicholson-Crotty, J., Grissom, J. A., & Nicholson-Crotty, S. (2011). Bureaucratic representation, distributional equity, and democratic values in the administration of public programs. *Journal of Politics*, *73*(2), 582–596. https://doi.org/10.1017/S0022381611000144

Paris, D. (2012). Culturally sustaining pedagogy: A needed change in stance, terminology, and practice. *Educational Researcher*, *41*(3), 93–97.

Solanki, S. M., & Xu, D. (2018). Looking beyond academic performance: The influence of instructor gender on student motivation in STEM fields. *American Educational Research Journal*, *55*(4), 801–835. https://doi.org/10.3102/0002831218759034

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*(6), 613–629. https://doi.org/10.1037/0003-066X.52.6.613

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*(5), 797–811. https://doi.org/10.1037/0022-3514.69.5.797

Wright, A., Gottfried, M. A., & Le, V. N. (2017). A kindergarten teacher like me: The role of student-teacher race in social-emotional development. *American Educational Research Journal*, *54*(1), 78S–101S. https://doi.org/10.3102/0002831216635733

Xu, D., & Li, Q. (2018). Gender achievement gaps among Chinese middle school students and the role of teachers' gender. *Economics of Education Review*, *67*, 82–93. https://doi.org/10.1016/j.econedurev.2018.10.002

Author

JENNIFER DARLING-ADUANA is an assistant professor of learning technologies in the Department of Learning Sciences at Georgia State University. She researches the equity implications of K–12 digital learning as well as the student–teacher and student–curriculum interactions in these settings that inadvertently contribute to (or can be used to mitigate) inequities in the classroom.