ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 21-399

The Effects of Teacher-Student Demographic Matching on Social-Emotional Learning

Ethan Scherer
Harvard University

Christopher Cleveland
Harvard University

Rebecca Ivester
Abt Associates

A growing body of research shows that students benefit when they are demographically similar to their teachers. However, less is known about how matching affects social-emotional development. We investigate the effect of teacher-student race and gender matching for middle school students in six charter management organizations. Using a student fixed effects strategy exploiting changes over time in the proportion of demographic matching in a school-grade, we estimate matching's effect on self-reports of interpersonal and intrapersonal social-emotional skills, test scores, and behavioral outcomes. We find improvements for Black and female students in interpersonal self-management and grit when they are matched to demographically similar teachers. We also find demographic matching leads to reductions in absences for Black students and improved math test scores for females. Our findings add to the emerging teacher diversity literature by showing its benefits for Black and female students during a critical stage of social-emotional development in their lives.

VERSION: May 2021

Suggested citation: Scherer, Ethan, Christopher Cleveland, and Rebecca Ivester. (2021). The Effects of Teacher-Student Demographic Matching on Social-Emotional Learning. (EdWorkingPaper: 21-399). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/3xq6-4k05

The Effects of Teacher-Student Demographic Matching on Social-Emotional Learning

Ethan Scherer
Center for Education Policy Research at Harvard University

Christopher Cleveland[1]
Harvard University

Rebecca Ivester
Abt Associates

Draft: April 2021

Abstract

A growing body of research shows that students benefit when they are demographically similar to their teachers. However, less is known about how matching affects social-emotional development. We investigate the effect of teacher-student race and gender matching for middle school students in six charter management organizations. Using a student fixed effects strategy exploiting changes over time in the proportion of demographic matching in a school-grade, we estimate matching's effect on self-reports of interpersonal and intrapersonal social-emotional skills, test scores, and behavioral outcomes. We find improvements for Black and female students in interpersonal self-management and grit when they are matched to demographically similar teachers. We also find demographic matching leads to reductions in absences for Black students and improved math test scores for females. Our findings add to the emerging teacher diversity literature by showing its benefits for Black and female students during a critical stage of social-emotional development in their lives.

* We thank the Walton Family Foundation for financial support and the Massachusetts Department of Elementary and Secondary Education and each of the charter management organizations for providing the administrative data for this work. Any errors of fact or interpretation are our own.

[1] The research reported here was supported, in whole or in part, by the Institute of Education Sciences, U.S. Department of Education, through grant R305B150010 to Harvard University. The opinions expressed are those of the authors and do not represent the views of the Institute or the U.S. Department of Education.

Introduction

Over the last decade, there has been increased focus on the importance of school-based outcomes other than test scores, especially social-emotional skills that predict both long-run academic performance and earnings (Chetty et al., 2010; Almlund et al., 2011; Heckman and Mosso, 2014). As momentum has gathered to collect these measures at scale, researchers have begun to descriptively explore how they change over time and differences between groups of students. Unlike test scores, these skills do not increase monotonically (West et al., 2020). Instead, for all students, self-reported skills like self-management and social awareness tend to decline during grades 6th through 8th and rebound in high school, identifying the middle school years as a key period of development and intervention. Furthermore, much like persistent achievement gaps that have been well documented in the past and continue to present themselves in recent analyses of race (de Brey et al., 2019) and gender (Reardon et al., 2019), researchers have also begun to describe significant differences in socio-emotional measures between races and genders (West et al., 2020). Fortunately, teachers are proving to be an effective lever for change in these less explored measures as they have been for traditional measures of academic success. Similar to prior research on highly effective teachers, as measured by student achievement (Chetty, Friedman and Rockoff, 2014), several recent studies have investigated a teacher's ability to improve social-emotional outcomes. These studies find that teachers who improve student skills and attitudes also improve student self-reported social-emotional measures (Blazer and Kraft, 2017) and help students obtain higher educational attainment (Jackson, 2018).

Given the importance of teachers in improving student outcomes, one proposed strategy to reduce the gaps between groups is to increase the racial/ethnic and gender diversity of the teaching workforce to reflect the diversity of the student population. The PK-12 student

population has grown increasingly diverse over the last two decades (de Brey et al., 2019), yet still 80 percent of US teachers are white, and 77 percent are female (US DOE, 2017).

The concept of increasing teacher workforce diversity has gained traction through a growing body of work that shows that demographically similar teachers can benefit students. Student-teacher demographic matching is associated with gains in student achievement (Dee, 2004; Clotfelter, Ladd and Vigdor, 2006; Goldhaber and Hansen, 2010; Egalite, Kisida and Winters, 2015; Harbatkin, 2021), reductions in absences and suspensions (Holt and Gershenson, 2015; Lindsay and Hart, 2017), and increases in the likelihood of high school graduation and college enrollment (Gershenson et al., 2019). While evidence continues to grow on the benefits of teacher demographic matching on student outcomes, the contexts and mechanisms that drive these changes continue to be debated. Furthermore, most of the research focuses on academic and behavioral outcomes rather than social-emotional skills that are increasingly gaining value in the modern workforce (Deming, 2017).

We address this gap by investigating demographic matching's effect on student social-emotional, academic, and behavioral outcomes in six diverse Boston charter management organizations. We test these hypotheses using a unique dataset of student social-emotional surveys in 5th to 8th grade and link these student surveys to rich administrative data for students and teachers. We then estimate the effects of increasing the proportion of demographically similar racial and gender teachers on student self-reports of interpersonal self-management, growth mindset, grit, and social awareness as well as both student academic and behavioral outcomes in the administrative data using a student fixed effects model. We find that even in the context of high expectation charter schools, there are positive effects for interpersonal and intrapersonal social-emotional measures as well as student behavior and test scores as a result of

racial and gender matching. These results are primarily driven by Black students and female students. Our evidence also finds that increasing racial staffing diversity improves social-emotional outcomes and math test scores for white students and indicates that increasing the diversity of the workforce is beneficial for all students in this context.

Our results contribute to the literature on student-teacher demographic matching in three ways. First, we are the first study to directly explore demographic matching on interpersonal skills, which are growing increasingly important in the labor market. Second, in a charter context that emphasizes teachers holding high expectations for all students, demographic matching still provides additional benefit to students. These results emphasize the value of continuing to diversify the teacher workforce, beyond workforce training and professional development opportunities. Third, we provide suggestive evidence that the primary mechanism driving these results is a role model effect: Teachers are serving as role models for students. This finding could provide important interim policy implications that should be tested, such as whether role models other than teachers (i.e., a high school or college student tutoring a middle school student) could provide similar results in the short run while other interventions to improve the diversity of the teacher labor market continue.

We structure the remainder of the paper in the following way: first, we begin by providing a brief description of our data, measures, and methods used in our analysis; next, we present our main results and an investigation of potential mechanisms; finally, we conclude with a discussion of our findings and their implications for policy efforts to increase teacher-student demographic matching.

Theory and Literature

We add to the to the social-emotional development literature by expanding upon the nascent work of how teaching teams contribute to changes in social-emotional measures while also reaffirming the discriminant validity of social-emotional survey measures. First, much of the teacher effectiveness research has focused on an individual teacher's contribution to student improvement using value-added models. However, particularly in the social-emotional domain, influence may not be confined to a single academic teacher. It is possible that a collaboration of teachers in both academic and non-academic domains (e.g., physical education teachers), school staff, and leadership work to improve student social-emotional skills. Jackson et al. (2020) show that school teams, including teachers, staff, and the administrators, that increase students' social-emotional skills not only predict reductions in high school behavioral infractions and increases in attendance and GPA, but also increases in four-year college going and college persistence. Our empirical strategy allows us to be inclusive of all full-time teachers, not limiting the sample to just the four primary academic subjects in middle school and isolates incremental effect of having one additional demographically matching teacher. Second, Jackson et al. (2020) demonstrate that we should not think of social-emotional skills as a single domain. They find evidence of discriminant validity between both interpersonal and intrapersonal social-emotional skills, like social health and academic motivation, respectively. For example, in horse-race regressions with standardized test scores and social-emotional skills, social health is predictive of reductions in absences and disciplinary incidence in 9^{th} grade, but academic motivation is not. Meanwhile academic motivation is predictive of post-secondary outcomes, but social health is not. Our data also allow us to uniquely explore both the inter and intrapersonal effects of demographic matching.

In addition, while consistent evidence on the student academic and behavioral benefits of teacher-student demographic matching continues to grow (Dee, 2004; Clotfelter, Ladd and Vigdor, 2006; Goldhaber and Hansen, 2010; Egalite, Kisida and Winters, 2015; Holt and Gershenson, 2015; Lindsay and Hart, 2017; Gershenson et al., 2019; Harbatkin, 2021), the mechanisms remain less well understood. As noted above, teachers play an important role in not only building cognitive skills, but also in developing socio-emotional skills. Demographic matching could operate through at least two mechanisms. The first can be broadly characterized as having a better cultural understanding. This connection either allows teachers to adapt the curriculum and provide more compelling or relevant examples (Villegas, Storm, and Lucas, 2012) or could reduce the likelihood of unconscious bias from negative stereotypes (Holt and Gershenson, 2015; Lindsay and Hart, 2017). In effect, this mechanism increases student academic and social-emotional skills through being a more effective teacher for demographically matched students. Thus, we expect the effects to be concentrated for matching demographic students because the "cultural understanding" would be more effective for matched students. We expect these results to mainly manifest through intrapersonal skills, like persistence and grit, rather than interpersonal effects because the teachers have deepened the interest of the students rather than the interaction between students (Gershenson et al., 2021).

The second way demographic matching could operate could be by allowing teachers to exert "role models" effects (Graham, 1987; Hess and Leal, 1997; Villegas, Storm, and Lucas, 2012). These role models could expand a student's information set on what is possible. For example, the demographically-matching teacher provides a concrete example to the student that someone "like them" can achieve academic success, augmenting the choice set over which the students optimize. The role model mechanism also suggests that additional teacher diversity

could benefit all students. For example, a teacher of color instructing a white student could help dispel stereotypes that the student has about other races or genders (Villegas, Storm, and Lucas, 2012). Relatedly, teachers who share similar demographic backgrounds as their students could have higher expectations for their students (Beady and Hansell, 1981; Gershenson et al., 2016). These higher expectations could also expand their choice sets. For example, by setting the expectation that all students can attend a post-secondary institution, teachers might encourage students to persist even if no one in their families has ever attended. We would expect these types of mechanisms to increase both intra and interpersonal measures since these mentors' own success stories likely required tenacity and building relationship with others (Gershenson et al., 2021).

Empirical work testing these theories has found evidence of both of them and that the mechanism could depend on the outcome (Egalite and Kisida, 2018; Gershenson et al., 2019). For example, Gershenson et al. (2019) develop a simple human capital model that can distinguish between potential mechanisms and put forth a series of testable assumptions of whether the role model or cultural competence explains the patterns they observe in the data. In their paper on long-run effects of Black student-teacher demographic matching in elementary schools, they find evidence of a stronger role model effect because of diminishing returns to the second Black teacher the student experiences, providing suggestive evidence that having a single advocate changes a student's information set into adulthood. However, Egalite and Kisida (2018) find evidence of both of these mechanisms using social-emotional measures. Understanding how these impacts vary based upon context and outcome is an important policy question as decision-makers invest their scarce resources.

Given the potential promise of demographic matching and the inherent difficulty of changing the supply chains for teachers, some have advocated for universal training to mimic some of these mechanisms. For example, as a short-term step, Egalite and Kisida (2018) advocate for educating teachers on tools to engage in culturally responsive teaching. However, implementation of trainings can vary in quality and without strong leadership from the principal and district, would likely not become a part of the school's culture. Another dimension that could be enhanced through training is setting high expectations for all students. High performing charter schools have been able to embed high expectations for all students into their culture. A key pillar of many high performing charters is that all students will attend a post-secondary institution and obtain professional careers. These expectations could help correct prior beliefs about the education production function (Egalite and Kisida, 2018; Gershenson et al., 2019). Furthermore, there is empirical evidence that some of these schools changed student post-secondary expectations and increased the number of students enrolling in a post-secondary education (Angrist et al., 2016; Coen, Nichols-Barrer, and Gleason, 2019; Davis and Heller, 2019). Testing if there is additional benefit to teacher-student demographic matching in this context expands the current literature by promoting an understanding of the generalizability of these findings across different contexts. It also provides a test case of how the expectations modeled in these charters influences the effects and mechanism of the demographic matching.

Data and Measures

Data Collection and Administrative Data

We collected data from 5th - 8th grade students attending Boston's public charter middle schools (BPS) during the 2015 to 2019 school years. Within these schools, we sampled all students from whom we received parental consent to participate and who attended school on the

data collection day. Students completed surveys assessing their social-emotional skills in their regular classrooms. We then merged these data with student-level administrative data—enrollment, attendance, suspensions, math and English language arts (ELA) test scores on the Massachusetts Comprehensive Assessment System (MCAS) and typical demographic information—provided by the Massachusetts Department of Elementary and Secondary Education (MADESE). MCAS scaled scores were standardized by grade, subject, and year by all students in Boston Public School to have mean zero and variance one. The baseline characteristics of the student sample are provided in Table 1. Of note, the majority of students in the sample are Black (47%) and/or recipients of free-or-reduced price lunch (92%).

Because we had the administrative data for all students, we compared the students who responded to the survey and the students who did not take the survey. We observe that on average these non-responding students performed worse academically and were more likely to absent or suspended, even though, on average, 78% of the students responded to the survey. To bound our estimates, we compare the results for the full administrative sample to those who completed the survey.[2]

We received administrative data from the teachers of students attending Boston's public charter middle schools during the 2015 to 2019 school years. We include all teachers in the participating schools. The characteristics of the teacher sample are provided in Table 2. Of note, the majority of teachers in the sample are White (72%) and/or female (77%). Next, we calculate the proportion of teachers who matched the student's demographics. Note that when calculating the match rate, the median student took about 8 classes in middle school. Many students not only

[2] We also explored if the samples varied by survey response year. For the years between 2016-2019, response rates varied between 78% and 81%, but was 73% in 2015 because some schools had trouble administering the survey. However, the demographics of the samples were not statistically different between 2015 and the other years.

took math, English, science and social studies, but also classes on art, foreign language or physical education. For students with more than 8 teachers, these numbers often included special education instructors, co-teachers, etc. The goal of including a wide range subjects is that social-emotional growth is not confined to the academic subjects. The anecdotes of the thoughtful physical education or inspirational music teacher do not fit into the academic boundaries. We provide information on the proportion of teachers who match a student's demographic characteristic by race and gender in Table 3. Although White students are not the majority of the sample, they are more likely to have a racial match with their teacher than students of other racial groups. Sixty-seven percent of White students have a racial match with a teacher in comparison to twenty-one percent of Black students.

Social-Emotional Survey Measures

Students in our sample completed online and paper surveys designed to measure their social-emotional skills in multiple areas. The survey consisted of scales on four constructs: grit, growth mindset, interpersonal self-management, and social awareness. These surveys were administered to students within their classrooms near the end of the academic year.

Intrapersonal Measures

Grit

Grit is a combination of passion and persistence. Individuals pursue goal with vigor and focus over an extended period of time. Lower measures of grit mean that you are easily discouraged or frequently led off track by new interests (Duckworth, et al., 2007). To measure student's grit, we used the Short Grit Scale (Farrington, 2013; Duckworth and Quinn, 2009). Students respond to four items (e.g., "I finish whatever I begin") on a five-point scale ranging

from “not like me at all” to “very much like me.” Overall grit is calculated by taking the average score for these four items and has an alpha of 0.76.

Growth Mindset

Growth Mindset is the belief that one’s abilities can grow with effort. Higher measures of growth mindset indicate one sees effort as necessary for success, embraces challenges, learns from criticism, and persists in the face of setbacks. To measure students’ growth mindset, we used the Implicit Theories of Intelligence Scale (Farrington et al., 2013; Blackwell et al., 2007). Students respond to four items (e.g., “challenging myself won’t make me any smarter) on a five-point scale ranging from “not at all true” to “completely true.” Overall growth mindset is calculated by reverse coding the items and taking the average score for these items and has an alpha of 0.70.

Interpersonal Measures

Interpersonal Self-Management

Self-Management is the ability to regulate one’s emotions, thoughts, and behaviors effectively in different situations. This includes managing stress, delaying gratification, motivating oneself, and setting and working toward personal and academic goals. To measure students’ interpersonal self-management, students completed a subset of the Impulsivity Scale for Children (Tsukayama, Duckworth, and Kim, 2013) which is comprised of nine items that measure students’ impulsivity as it relates to behavior, attention, and emotions. Items asked students to indicate how often, on a five-point scale ranging from “almost never” to “almost always”, in the past 30 days they exhibited a specific set of behaviors. Four items assessed interpersonal self-control (e.g., “I interrupted other students while they were talking”). We

calculated their overall interpersonal self-management score by reverse-coding and averaging these four items and has an alpha of 0.79.

Social Awareness

Social Awareness is the ability to take the perspective of and empathize with others from diverse backgrounds and cultures, to understand social and ethical norms for behavior, and to recognize family, school, and community resources and supports. To measure students' social awareness, we adapted from AIR and CASEL (2013) student self-report of social and emotional competencies. Students respond to eight items based upon their experiences during the last school year (e.g., "I listened carefully to other people's points of view") on a five-point scale ranging from "almost never" to "almost all the time." We calculated their overall social awareness score by averaging these eight items and has an alpha of 0.86.

Empirical Strategy

Our identification strategy is similar to Lindsay and Hart (2017) that leverages the fact that the proportion of demographically matching teachers a student experiences will vary across each grade, but controls for any time-invariant observable and unobservable grade and time characteristics that are related to our outcomes of interest. The model specification is the following:

$$Y_{igst} = B_0 + \beta_1 PropTeachRace_{igst} + \beta_2 PropTeachGend_{igst} + \beta_3 PropStudGend_{igst} + \beta_4 PropStudRace_{igst} + \sigma_g + \omega_t + \tau_i + \epsilon_{igst}$$

Y_{gsit} represents student i's outcome, in grade g in school s in period t. The predictors of interest are $PropTeachRace_{igst}$ and $PropTeachGender_{igst}$ representing student i's proportion of

teachers who share student i's race or gender in grade g at school s at time *t*.[3] We also include $PropStudRace_{igst}$ and $PropStudGend_{igst}$ representing student *i*'s proportion of students in the school-grade who share student i's race or gender at time *t*. We include student fixed effects (τ_i), grade fixed effects σ_g as well as time fixed effects ω_t. The coefficient on $PropTeachRace_{igst}$ and $PropTeachRace_{igst}$ is generated from the variation within a given student, over time. The proportion of demographically matching teachers varies by year allowing us to estimate the effect of having more or fewer demographically teachers across time. Standard errors are clustered at the school-grade level. For ease of interpretation, a one-unit change is equivalent to a 10 percentage point change in the student-teacher demographic matching rate. We include all teachers who worked with the student, as mention above. In our sample, the median student had 8 teachers. Thus, a 10 percentage point change, on average, is exposure to one additional teacher in these schools.

Results

Social-Emotional Outcomes

We first examine how the proportion of teacher and student race and gender match are associated with social-emotional learning outcomes. Table 4 reports the outcome in the rows and the coefficients of interest, proportion of same teacher race and gender are on column (2) and (4) respectively, with the standard errors in (3) and (5) respectively. For each outcome, the first row reports the overall effects. Then we report the effects for several subgroups by subsetting the analysis for the specific population of students and running the student fixed effects regression again.

[3] We model both the linear and quadratic versions of the main predictors and present the quadratic results in the Appendix. For females there are diminishing returns on teaching matching for grit and interpersonal management, but little evidence in other cases.

Overall matching on the same race increases interpersonal self-management. A 10 percentage point increase in the proportion of student-teacher matching on race in the school-grade, on average about one additional teacher, increases interpersonal self-management by 0.02 standard deviations units. In our sample, across all students and grades, students tend to decline about 0.016 standard deviations each year in interpersonal self-management. Thus, the magnitude of these effects is quite large, about a full year of change in middle school. Overall, the remaining constructs cannot be differentiated from zero for race. However, for gender matching, overall, we observe a 10 percentage point increase in similar gender teacher-student matching results in improvements in of 0.01for interpersonal self-management and grit. Declines in grit are much steeper for our sample than interpersonal self-management, where on average students decline 0.053 standard deviation units each year in middle school. While not as large as the overall effects on race, these are still practically meaningful.

However, these overall effects mask important differences in teacher-student race matching in particular. This is mainly due to the fact that among the white students in our schools, there is a benefit to increased staff diversity. For all the constructs, the coefficient on proportion of same race for white students is negative, and marginally significant for grit. Thus, among white students, a 10 percentage point *decrease* in student-teacher matching *increases* student social-emotional skills. Even though white students represent a small proportion of our sample, these negative effects mask the overall positive effects of increasing diversity due to the negative effect of same race matching for white students. Among Black students, a 10 percentage point increase in student-teacher matching increases interpersonal self-management and grit by 0.02 standard deviations and social awareness at the 0.10 level of significance. Thus, there are clear social-emotion benefits when the racial diversity of the staff increases. For gender, we

observe that the increases in student-teacher gender matching effects in interpersonal self-management and grit were driven by females. Thus, among female students, for a 10 percentage point increase in student-teacher matching increases interpersonal self-management and grit by 0.02 standard deviations. Broadly we observe that matching on race or gender tends to improve student self-reported persistence and social interactions, like not interrupting other students when they are speaking.

Academic and Behavioral Outcomes

We next examine how the proportion of teacher and student race and gender match are associated with academic and behavioral outcomes (Table 5). Overall, we observe no significant benefit of race or gender matching on test scores or behavioral outcomes, like ever being suspended or the log of number of school days absent.[4] While the coefficients on the log of the number of days absent suggests that matching reduced the absences, these effects are not statistically significant, and the magnitude of the overall effect is a one percent reduction.

When we explore different subgroups of students, many of the patterns we observed in the social-emotional outcomes are similar for math test scores and absences. For math test scores, white students improve 0.02 standard deviations at the $p<0.10$ level, as the staff racial demographics differ from the student. However, particularly for absences, Black students tend to attend more school when there is demographic matching, and white students tend to attend when there is greater diversity of the staff, but the white coefficient is not statistically significant. Our results for demographic match show stronger relationships for males and female. For females, math test score improves when they have more teachers who share their gender. For a 10 percentage point increase in the proportion of female matching increases math test scores by 0.02

[4] We also estimated these models using the natural log of 1+ number of suspensions and these are directionally similar, so we just report the limited probability model on ever being suspended.

standard deviations. For males an increase of about one male teacher reduced ever being suspended by 1% and reduced number of days absent by one percent, though the former is only significant at the 0.10 level and the latter is not significant. However, for males, a 10 percentage point increase in male teachers reduces their math test scores by 0.02 standard deviations. Thus, increasing the exposure of the proportion of female teachers for females improves their academics, and for males improves their behavior.

In the descriptive statistics, we noted that both students that were lower performing on test scores and had more suspensions and absences were not in our social-emotional analytic sample. To understand how the administrative effects differed for these students, we also analyzed the sample where we had complete administrative data. Table 6 shows that including these students maintains our test scores results and improves the absence patterns in our main analytics sample. For a 10 percentage point increase in the proportion of matching teachers for Black students there was a 2 percentage decrease in number of days absent. In contrast, the effects of increased staff diversity statistically decrease absences for white students. In summary, the effect of including the students who did not complete the social-emotional survey was generally to strengthen our administrative data findings, indicating our social-emotional effects could be an underestimate.

Table 7 explores the robustness of our administrative data results using different specifications of the models. We focus on Math test scores and absences, since these were outcomes where the effects were significant at the 0.10 level. The top row in each subgroup provides the baseline estimates from Table 6 to ease comparisons. The second row in each subgroup adds school fixed effects to each model. The third row adds characteristics of the student's school-grade including the proportion of students receiving free/reduced price lunch,

students with an Individualized Education Program, English language learners, and the average prior year test scores in math and ELA. The fourth row adds teacher characteristics of the grade, like the proportion of teachers for each race and gender, as well as average experience and experience squared of the school-grade team. These results are similar in direction and magnitude as the baseline results indicating that our administrative results are not sensitive to these changes.

Discussion

Using a unique longitudinal dataset, we estimate a student fixed effects model that identifies the effect of increasing the proportion of demographically matching teachers on both social-emotional survey measures and administrative data outcomes in the context of high expectation charter middle schools. For Black students, we find practically large effects on both intra and interpersonal measures that also reduce absences by 0.08 percentage points. In addition, we find similar size effects on both intra and interpersonal measures for female students that increase math test scores by 0.02 standard deviations. Interestingly, there is suggestive evidence that the small percentage of white students in these schools benefit from non-matching teachers, with marginally significant improvements in grit and behavioral engagement measures like absences.

These results are important for three reasons. First, they broaden our understanding of how teacher and student demographic matching affects specific student skills. In our schools, demographic matching improves both personal motivation and how students interact with one another. These types of skills are becoming more important in our economy (Deming, 2017) and if these effects persist, as has been found in Gershenson et al. (2019), they could provide significant benefits in the long run. Middle school is a particularly tumultuous time, where self-

management and social awareness tends to decline for all students (West et al., 2020), so identifying interventions that assist during this time could be helpful. Second, we build on prior evidence (Gershenson et al., 2019) empirically identifying demographic matching's mechanism. We interpret the positive effect on white students and equal size magnitude on the inter and intrapersonal measures for Black and female students as suggestive of the "role model" mechanism. Because our model focuses on the proportion of similar demographic teachers for all full-time teachers, including non-core subjects, these results provide some interesting future testable assumptions. For example, could community group members or tutors also play an important mentorship role? While we do not argue these individuals should substitute for the continued effort to diversify the teaching workforce, the labor market has been slow to change and these individuals could provide short-run policy solutions to move the work forward.

Finally, the effects that we observe on female students suggest that policymakers should consider new issues that could arise as they move forward in diversifying the workforce. In our schools, the majority of the staff are women, like the rest of the nation. However, we continue to find strong benefits for female students of having a female teacher. As noted in Appendix A, there are diminishing returns to female gender match with the maximum occurring around 60-70% of the student's teachers. Additionally, even though we find positive implications for students on increasing the racial diversity of the teacher workforce, Black students match their teacher 18% of the time in our sample. Thus, as policymakers continue to design ways to change the workforce, and these rates move higher, the curvature could change. Thus, policymakers should consider potential unintended consequences of focusing on a particular group. For example, much of the prior research has found positive effects for Black male students to have a Black male teacher (Egalite and Kisida, 2018; Gershenson et al., 2019). While these match rates

remain relatively low in most systems, implications for females during structural change should continue to be considered.

A few limitations apply to this work. First, compared to prior evidence, we have a relatively smaller sample size and focus on a specific set of charter schools that have historically been shown to be effective at increasing academic achievement and college enrollment. These differences limit the generalizability of our findings even to other charter schools and it could be helpful for others to replicate our work in other locations and among other charter schools. Second, it could be that the positive effects that we observe are not the effect of race/gender-matching but could be indicative of differences in teacher quality. For reasons discussed earlier, estimating social-emotional value-added for a particular teacher could incorrectly attribute social-emotional growth to that teacher. Furthermore, social-emotional value-added is still in its infancy and preliminary work has found that the estimates tend to be less stable than test scores (Loeb et al., 2019) and often are not highly correlated with test score value-added (Jackson, 2018). Third, while the demographic matching literature has grown over the last couple decades, it often relies on quasi-experimental methods using large secondary datasets, with the notable exception of Dee (2004) who conducted an experiment. Future research should confirm our findings using experimental data in multiple locations and with varied agency types.

References

Almlund, M., Duckworth, A. L., Heckman, J., & Kautz, T. (2011). Personality psychology and economics. In *Handbook of the Economics of Education* (Vol. 4, pp. 1-181). Elsevier. doi: 10.1016/B978-0-444-53444-6.00001-8

Angrist, J., Hull, P., Pathak, P., & Walters, C. (2016). Interpreting tests of school VAM validity. *American Economic Review*, *106*(5), 388-92.

Adapted from AIR and CASEL (2013) Student self-report of social and emotional competencies

Black, D. S., Sussman, S., Johnson, C. A., and Milam, J. (2012). Psychometric assessment of the mindful attention awareness scale (MAAS) among Chinese adolescents. *Assessment*, 19(1), 42-52.

Beady Jr, C. H., & Hansell, S. (1981). Teacher race and expectations for student achievement. *American Educational Research Journal*, *18*(2), 191-206.

Blazar D, Kraft MA. Teacher and teaching effects on students' attitudes and behaviors. *Educational Evaluation and Policy Analysis*. 2017;39 (1) :146-170.

de Brey, C., Musu, L., McFarland, J., Wilkinson-Flicker, S., Diliberti, M., Zhang, A., ... & Wang, X. (2019). Status and Trends in the Education of Racial and Ethnic Groups 2018. NCES 2019-038. *National Center for Education Statistics*. Retrieved April 27, 2021 from https://nces.ed.gov/ pubsearch/.

Chetty, R, Friedman, J. N, Hilger, N, Saez, E, Schanzenbach, D. W, and Yagan, D. (2011). How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star. *The Quarterly Journal of Economics*, 126(4), 1593-1660.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review*, *104*(9), 2593-2632.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American economic review*, *104*(9), 2633-79.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2006). Teacher-student matching and the assessment of teacher effectiveness. *Journal of human Resources*, *41*(4), 778-820.

Davis, M., & Heller, B. (2019). No Excuses charter schools and college enrollment: New evidence from a high school network in Chicago. *Education Finance and Policy*, *14*(3), 414-440.

Dee, T. S. (2004). The Race Connection. Education next, 4(2), *Education Next*, 2004-04-01, Vol.4 (2).

Deming, D. J. (2017). The Growing Importance of Social Skills in the Labor Market. *The Quarterly Journal of Economics*, 132(4), 1593-1640.

Egalite, A.J., Kisida, B. and Winters, M.A. (2015). Representation in the classroom: The effect of own-race teachers on student achievement. *Economics of Education Review*, 45(1), 44-52.

Egalite, A. J., & Kisida, B. (2018). The effects of teacher match on students' academic perceptions and attitudes. *Educational Evaluation and Policy Analysis*, *40*(1), 59-81.

Farrington, C., Levenstein, R., & Nagaoka, J. (2013). " Becoming Effective Learners" Survey Development Project. *Society for Research on Educational Effectiveness*.

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child development*, *78*(1), 246-263.

Gehlbach, H., Brown, S. W., Ioannou, A., Boyer, M. A., Hudson, N., Niv-Solomon, A., ... & Janik, L. (2008). Increasing interest in social studies: Social perspective taking and self-efficacy in stimulating simulations. *Contemporary Educational Psychology*, *33*(4), 894-914.

Gershenson, S., & Holt, S. B. (2015). Gender gaps in high school students' homework time. *Educational Researcher*, *44*(8), 432-441. DOI: 10.3102/0013189X1561612

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.. DOI: 10.1016/j.econedurev.2016.03.002

Gershenson, S., Hart, C., Hyman, J., Lindsay, C., & Papageorge, N. W. The Long-Run Impacts of Same-Race Teachers. (EdWorkingPaper: 19-43). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/9419-nw68

Gershenson, S., Lindsay, C., and Hansen, M. (2021). Teacher Diversity and Student Success: Why Racial Representation Matters in The Classroom. Harvard Education Press.

Goldhaber, D., & Hansen, M. (2010). Race, gender, and teacher testing: How informative a tool is teacher licensure testing?. *American Educational Research Journal*, *47*(1), 218-251.

Goldhaber, D., Quince, V., & Theobald, R. (2018). Has it always been this way? Tracing the evolution of teacher quality gaps in US public schools. *American Educational Research Journal*, *55*(1), 171-201.

Graham, J. G. (1987). English Language Proficiency and the Prediction of Academic Success. *TESOL Quarterly*, 21(3), 505-521.

Harbatkin, E. (2021). Does student-teacher race match affect course grades?. *Economics of Education Review*, *81*, 102081.

Heckman, J. J., & Mosso, S. (2014). The economics of human development and social mobility. *Annu. Rev. Econ.*, *6*(1), 689-733.

Hess, F. M., & Leal, D. L. (1997). Minority teachers, minority students, and college matriculation: A new look at the role-modeling hypothesis. *Policy Studies Journal*, *25*(2), 235-248.

Holt, S. B., and Gershenson, S. 2017. The impact of teacher demographic representation on student attendance and suspensions. In Press, *Policy Studies Journal*. DOI: 10.1111/psj.12229.

Jackson, C. K. (2018). What Do Test Scores Miss? The Importance of Teacher Effects on Non–Test Score Outcomes. *The Journal of Political Economy*, 126(5), 2072-2107.

Jackson, C.L, Porter S. K., Easton, J.Q., Blanchard, A. "School Effects on Socio-emotional Development, School-Based Arrests, and Educational Attainment" *American Economic Review: Insights* (2020).

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

Kraft, M. A., Grace, S. (2016). Teaching for tomorrow's economy? Teacher effects on complex cognitive skills and social-emotional competencies (Working paper). Providence, RI: Brown University. Retrieved from http://scholar.harvard.edu/files/mkraft/files/teaching_for_tomorrows_economy_-_final_public.pdf

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.

Loeb, S., Christian, M. S., Hough, H., Meyer, R. H., Rice, A. B., and West, M. R. (2019). School differences in social–emotional learning gains: Findings from the first large-scale panel survey of students. *Journal of Educational and Behavioral Statistics*, *44*(5), 507-542.

Papageorge, N. W., Gershenson, S., & Kang, K. M. (2020). Teacher expectations matter. *Review of Economics and Statistics*, *102*(2), 234-251.

U.S. Department of Education. (2019, February). *Spotlight A: Characteristics of Public School Teachers by Race/Ethnicity*. Retrieved from https://nces.ed.gov/programs/raceindicators/spotlight_a.asp

Villegas, A. M., Strom, K., & Lucas, T. (2012). Closing the racial/ethnic gap between students of color and their teachers: An elusive goal. *Equity & Excellence in Education*, *45*(2), 283-301.

Coen, T., Nichols-Barrer, I., & Gleason, P. (2019). Long-Term Impacts of KIPP Middle Schools on College Enrollment and Early College Persistence. *Mathematica*.

West, M. R., Pier, L., Fricke, H., Hough, H., Loeb, S., Meyer, R. H., & Rice, A. B. (2020). Trends in Student Social-Emotional Learning: Evidence From the First Large-Scale Panel Student Survey. *Educational Evaluation and Policy Analysis*, 42(2), 279–303. https://doi.org/10.3102/0162373720912236

Table 1: Descriptive Statistics of Analytic and Administrative Sample

	Survey Sample	Non-Survey Sample	Difference
Student Characteristics			
Male	0.494	0.547	-0.053***
	(0.500)	(0.498)	(0.010)
White	0.090	0.070	0.020**
	(0.287)	(0.255)	(0.009)
Latinx	0.285	0.262	0.023**
	(0.451)	(0.440)	(0.011)
Black	0.471	0.530	-0.058***
	(0.499)	(0.499)	(0.017)
Other	0.154	0.138	0.015
	(0.361)	(0.345)	(0.010)
Individual Education Plan	0.231	0.314	-0.083***
	(0.421)	(0.464)	(0.013)
English Language Learner	0.425	0.436	-0.011
	(0.494)	(0.496)	(0.016)
Free/Reduced Price Lunch	0.918	0.933	-0.015**
	(0.274)	(0.250)	(0.007)
Student Administrative Data Outcomes			
Math Test Scores	0.204	-0.031	0.235***
	(0.875)	(0.918)	(0.036)
ELA Test Scores	0.087	-0.128	0.215***
	(0.888)	(0.926)	(0.031)
Ever Suspended	0.161	0.267	-0.106***
	(0.368)	(0.442)	(0.012)
Percent Days Absent	3.904	6.264	-2.359***
	(3.861)	(6.974)	(0.220)
Unique Students	4490	931	5421
N	11,758	3,280	15,038

Notes: Students in the Survey Sample answered all survey items and appeared in our data at least twice. The Non-Survey Sample constitute all other students who appeared at least twice during and we have absence/suspension data. The Other demographic category includes Asian, American Indian, Alaskan, Native Hawaiian or Pacific Islander and multiple races. Standard deviations are shown in parentheses for the Survey and Non-Survey column. ELA = English Language Arts. + $p<0.10$,*$p<0.05$,**$p<0.01$ refer to mean-difference using school-grade clustered standard errors between surveyed students and non-surveyed students.

Table 2: Descriptive Statistics of the Teacher Sample

	Mean	SD
Female	0.77	0.42
White	0.72	0.45
Black	0.18	0.39
Other	0.06	0.24
Unique Teachers	494	

Notes: The Other demographic category includes Asian, American Indian, Alaskan, Native Hawaiian or Pacific Islander and multiple races.

Table 3: Sample Mean of Key Independent Variables, by Student Characteristics

	Same Race	Same Gender
All	0.19	0.50
White	0.67	0.51
Black	0.21	0.50
Female	0.18	0.67
Male	0.19	0.33

Table 4: Effects of Teacher/Student Demographic Match on Social-Emotional Outcomes using Fixed Effects Model

		Prop Same Race Tch		*Prop Same Gender Tch*		*Observations*	*Adjusted R-squared*
		B	*SE*	*B*	*SE*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Interpersonal-Self-Management	Overall	0.02*	(0.007)	0.01*	(0.005)	11758	0.514
	White	-0.01	(0.014)	-0.00	(0.010)	1029	0.484
	Black	0.02*	(0.008)	0.01+	(0.006)	5438	0.510
	Female	0.01	(0.009)	0.02*	(0.006)	5947	0.527
	Male	0.02*	(0.008)	<0.01	(0.007)	5801	0.498
Growth Mindset	Overall	<0.01	(0.007)	<0.01	(0.006)	11758	0.425
	White	-0.02	(0.019)	0.01	(0.009)	1029	0.427
	Black	0.01	(0.007)	<0.01	(0.008)	5438	0.420
	Female	<0.01	(0.011)	0.01	(0.009)	5947	0.439
	Male	0.01	(0.010)	<-0.01	(0.009)	5801	0.410
Grit	Overall	0.01	(0.007)	0.01**	(0.004)	11758	0.472
	White	-0.03+	(0.018)	0.01	(0.022)	1029	0.430
	Black	0.02*	(0.008)	0.02**	(0.006)	5438	0.456
	Female	0.01	(0.010)	0.02*	(0.008)	5947	0.475
	Male	<0.01	(0.011)	<0.01	(0.007)	5801	0.470
Social Awareness	Overall	0.01	(0.007)	<0.01	(0.006)	11758	0.462
	White	-0.01	(0.012)	-0.01	(0.011)	1029	0.448
	Black	0.02+	(0.009)	0.01	(0.007)	5438	0.458
	Female	0.01	(0.009)	<0.01	(0.007)	5947	0.498
	Male	0.01	(0.011)	<-0.01	(0.013)	5801	0.423

Notes: Standard errors are reported in parentheses and clustered at the school-grade level. Samples are restricted to students with social-emotional survey. All models also include student fixed effects. All survey measures are standardized by grade-level within the sample of students who completed the survey.
+ p<0.10,*p<0.05,**p<0.01

Table 5: Effects of Teacher/Student Demographic Match on Test Scores and Observed Behavior Outcomes using Fixed Effects Model for Students with Social-Emotional Survey Data

		Prop Same Race Tch		*Prop Same Gender Tch*		*Observations*	*Adjusted R-squared*
		B	*SE*	*B*	*SE*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Math Test Score	Overall	0.01	(0.007)	<0.01	(0.002)	11579	0.827
	White	-0.02+	(0.011)	<0.01	(0.007)	1016	0.816
	Black	0.01	(0.011)	<-0.01	(0.003)	5348	0.827
	Female	<0.01	(0.007)	0.02**	(0.007)	5859	0.829
	Male	<0.01	(0.007)	-0.02*	(0.008)	5710	0.827
English Language Arts Test Scores	Overall	<0.01	(0.006)	-0.00	(0.003)	11605	0.732
	White	0.01	(0.017)	-0.01+	(0.006)	1022	0.733
	Black	<0.01	(0.010)	<-0.01	(0.005)	5354	0.728
	Female	-0.00	(0.007)	<0.01	(0.006)	5881	0.727
	Male	<0.01	(0.008)	<-0.01	(0.007)	5716	0.719
Ever Suspended	Overall	<0.01	(0.003)	<-0.01	(0.002)	11758	0.361
	White	-0.00	(0.007)	-0.01	(0.005)	1029	0.355
	Black	0.01	(0.005)	<0.01	(0.003)	5438	0.366
	Female	<0.01	(0.004)	<0.01	(0.002)	5947	0.326
	Male	0.01	(0.005)	-0.01+	(0.005)	5801	0.353
Log Number of Days Absent	Overall	-0.01	(0.006)	-0.00	(0.003)	11758	0.624
	White	0.01	(0.015)	0.01	(0.011)	1029	0.622
	Black	-0.01*	(0.006)	<-0.01	(0.005)	5438	0.609
	Female	-0.01	(0.010)	<0.01	(0.006)	5947	0.629
	Male	-0.01+	(0.007)	-0.01	(0.006)	5801	0.619

Notes: Standard errors are reported in parentheses and clustered at the school-grade level. Samples are restricted to students with outcome data. All models also include student, year, and grade fixed effects. We standardize the test scores across all students in Boston Public Schools.

+ p<0.10,*p<0.05,**p<0.01

Table 6: Effects of Teacher/Student Demographic Match on Test Scores and Observed Behavior Outcomes using Fixed Effects Model Including All Students

		Prop Same Race Tch		*Prop Same Gender Tch*		*Observations*	*Adjusted R-squared*
		B	*SE*	*B*	*SE*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Math Test Score	Overall	<0.01	(0.006)	<0.01	(0.002)	14735	0.825
	White	-0.02+	(0.011)	0.01	(0.009)	1240	0.807
	Black	0.01	(0.010)	<-0.01	(0.003)	6987	0.825
	Female	<0.01	(0.007)	0.02*	(0.007)	7293	0.831
	Male	-0.00	(0.007)	-0.02*	(0.008)	7430	0.820
English Language Arts Test Scores	Overall	0.00	(0.005)	<-0.01	(0.003)	14761	0.732
	White	0.01	(0.015)	-0.01	(0.005)	1248	0.724
	Black	<0.01	(0.008)	-0.01+	(0.004)	6986	0.728
	Female	<0.01	(0.006)	<-0.01	(0.006)	7325	0.727
	Male	0.01	(0.007)	<-0.01	(0.007)	7426	0.717
Ever Suspended	Overall	<0.01	(0.003)	<-0.01	(0.002)	15038	0.390
	White	<0.01	(0.007)	-0.01	(0.005)	1257	0.280
	Black	<0.01	(0.004)	<0.01	(0.003)	7157	0.390
	Female	<-0.01	(0.003)	<-0.01	(0.003)	7433	0.335
	Male	<0.01	(0.004)	-0.01	(0.005)	7593	0.390
Log Number of Days Absent	Overall	-0.01	(0.007)	-0.00	(0.003)	15038	0.637
	White	0.02+	(0.014)	<0.01	(0.010)	1257	0.617
	Black	-0.02*	(0.008)	-0.01	(0.004)	7157	0.624
	Female	-0.01	(0.010)	<0.01	(0.007)	7433	0.642
	Male	-0.01	(0.008)	-0.01	(0.006)	7593	0.633

Notes: Standard errors are reported in parentheses and clustered at the school-grade level. Samples are restricted to students with outcome data. All models also include student, year, and grade fixed effects. We standardize the test scores across all students in Boston Public Schools.
+ p<0.10,*p<0.05,**p<0.01

Table 7: Specification Checks on Full Administrative Sample

			Prop Same Race Tch		*Prop Same Gender Tch*		*Obs*	*Adjusted R-squared*
			B	*SE*	*B*	*SE*		
Math Test Scores	White	Baseline	-0.02+	(0.011)	0.01	(0.009)	1240	0.807
		Add School FE	-0.03*	(0.011)	0.01	(0.008)	1240	0.815
		Add School-Grade Demo	-0.02	(0.011)	0.01	(0.010)	1240	0.809
		Add School-Grade Tch	-0.01	(0.011)	0.01	(0.008)	1240	0.813
	Female	Baseline	<0.01	(0.007)	0.02*	(0.007)	7293	0.831
		Add School FE	<0.01	(0.007)	0.02**	(0.006)	7293	0.836
		Add School-Grade Demo	<0.01	(0.006)	0.01*	(0.007)	7293	0.832
		Add School-Grade Tch	<-0.01	(0.006)	0.01+	(0.007)	7293	0.834
	Males	Baseline	<-0.01	(0.007)	-0.02*	(0.008)	7430	0.820
		Add School FE	-0.01	(0.006)	-0.02*	(0.008)	7430	0.824
		Add School-Grade Demo	<-0.01	(0.006)	-0.01+	(0.008)	7430	0.822
		Add School-Grade Tch	-0.01	(0.006)	-0.01	(0.008)	7430	0.823
Log Number of Days Absent	Overall	Baseline	-0.01	(0.007)	<-0.01	(0.003)	15038	0.637
		Add School FE	-0.01+	(0.007)	<-0.01	(0.003)	15038	0.640
		Add School-Grade Demo	-0.01	(0.007)	<-0.01	(0.003)	15038	0.638
		Add School-Grade Tch	-0.01	(0.007)	<-0.01	(0.003)	15038	0.639
	White	Baseline	0.02+	(0.014)	<0.01	(0.010)	1257	0.617
		Add School FE	0.03*	(0.013)	0.01	(0.009)	1257	0.622
		Add School-Grade Demo	0.02+	(0.013)	<0.01	(0.010)	1257	0.618
		Add School-Grade Tch	0.02	(0.013)	0.01	(0.009)	1257	0.630
	Black	Baseline	-0.02*	(0.008)	-0.01	(0.004)	7157	0.624
		Add School FE	-0.02**	(0.008)	<-0.01	(0.004)	7157	0.628
		Add School-Grade Demo	-0.02*	(0.008)	<-0.01	(0.004)	7157	0.625
		Add School-Grade Tch	-0.01	(0.008)	<-0.01	(0.004)	7157	0.626
	Female	Baseline	-0.01	(0.010)	<0.01	(0.007)	7433	0.642
		Add School FE	-0.02+	(0.010)	<0.01	(0.007)	7433	0.644
		Add School-Grade Demo	-0.01	(0.010)	<0.01	(0.006)	7433	0.642
		Add School-Grade Tch	-0.01	(0.010)	<0.01	(0.006)	7433	0.643
	Males	Baseline	-0.01	(0.008)	-0.01	(0.006)	7593	0.633
		Add School FE	-0.01+	(0.007)	-0.01+	(0.006)	7593	0.636
		Add School-Grade Demo	-0.01	(0.008)	-0.01+	(0.006)	7593	0.633
		Add School-Grade Tch	-0.01	(0.007)	-0.01*	(0.006)	7593	0.635

Notes: All models also include student, year, and grade fixed effects as well as the proportion of students' own race and gender in the school-grade. The school-grade demographics include the proportion of free/reduced price lunch, individual education plans and English language learners' students in the class, and the average prior year math and ELA performance. Missing scores are imputed to the mean. The school-grade teacher models include the proportion of teachers race and gender in the school-grade as well as the average experience and experienced squared for the school-grade teaching team. Standard errors are reported in parentheses and clustered at the school-grade level. Samples are restricted to students with outcome data. We standardize the test scores across all students in Boston Public Schools. + p<0.10,*p<0.05,**p<0.01

Table A1: Effects of Non-Linear Teacher/Student Demographic Match on Social-Emotional Outcomes using Fixed Effects Model

		Prop Same Race Tch		*Prop Same Race Tch^2*		*Prop Same Gender Tch*		*Prop Same Gender Tch^2*		*Obs*	*Adjusted R-squared*
		B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*		
Interpersonal-Self-Mgmt	Overall	0.03*	(0.012)	-0.00	(0.002)	0.03	(0.019)	<-0.01	(0.002)	11758	0.514
	White	-0.05	(0.050)	<0.01	(0.004)	-0.05	(0.033)	<0.01	(0.003)	1029	0.483
	Black	0.02	(0.026)	<-0.01	(0.004)	0.02	(0.019)	<-0.01	(0.002)	5438	0.510
	Females	<0.01	(0.013)	<0.01	(0.002)	0.10**	(0.027)	-0.01**	(0.002)	5947	0.528
	Males	0.05*	(0.019)	-0.01*	(0.002)	0.02	(0.022)	<-0.01	(0.002)	5801	0.499
Growth Mindset	Overall	0.03+	(0.015)	<-0.01	(0.002)	<0.01	(0.021)	<0.01	(0.002)	11758	0.425
	White	0.03	(0.060)	<-0.01	(0.005)	0.03	(0.035)	<-0.01	(0.003)	1029	0.426
	Black	0.03	(0.024)	<-0.01	(0.004)	-0.02	(0.024)	<0.01	(0.002)	5438	0.420
	Females	0.03	(0.021)	<-0.01	(0.003)	0.02	(0.036)	<-0.01	(0.003)	5947	0.439
	Males	0.02	(0.023)	<-0.01	(0.003)	0.02	(0.028)	<-0.01	(0.003)	5801	0.410
Grit	Overall	0.03+	(0.014)	<-0.01+	(0.002)	0.04	(0.022)	<-0.01	(0.002)	11758	0.472
	White	0.03	(0.070)	<-0.01	(0.006)	0.04	(0.032)	<-0.01	(0.003)	1029	0.429
	Black	0.05+	(0.024)	-0.01	(0.005)	0.02	(0.027)	<-0.01	(0.003)	5438	0.456
	Females	0.03+	(0.019)	-0.00+	(0.002)	0.12**	(0.040)	-0.01*	(0.003)	5947	0.477
	Males	0.01	(0.023)	<-0.01	(0.003)	0.04	(0.025)	<-0.01	(0.003)	5801	0.470
Social Awareness	Overall	0.03+	(0.014)	<-0.01	(0.002)	<0.01	(0.031)	<-0.01	(0.003)	11758	0.462
	White	-0.04	(0.052)	<0.01	(0.004)	-0.07	(0.043)	0.01	(0.004)	1029	0.447
	Black	0.05	(0.029)	<-0.01	(0.005)	-0.01	(0.032)	<0.01	(0.003)	5438	0.458
	Females	0.03	(0.018)	<-0.01	(0.003)	0.01	(0.033)	<-0.01	(0.003)	5947	0.498
	Males	0.02	(0.022)	<-0.01	(0.003)	0.01	(0.040)	<-0.01	(0.004)	5801	0.423

Notes: Standard errors are reported in parentheses and clustered at the school-grade level. Samples are restricted to students with outcome data. All models also include student, year, and grade fixed effects. We standardize the test scores across all students in Boston Public Schools.
+ p<0.10,*p<0.05,**p<0.01

Table A2: Effects of Non-Linear Teacher/Student Demographic Match on Test Scores and Observed Behavior Outcomes

		Prop Same Race Tch		*Prop Same Race Tch^2*		*Prop Same Gender Tch*		*Prop Same Gender Tch^2*		*Obs*	*Adjusted R-squared*
		B	*SE*	*B*	*SE*	*B*	*SE*	*B*	*SE*		
Math Test Scores	Overall	0.01	(0.014)	<-0.01	(0.002)	-0.05*	(0.020)	<0.01*	(0.002)	11579	0.828
	White	-0.06	(0.039)	<0.01	(0.003)	-0.05+	(0.030)	0.01+	(0.003)	1016	0.817
	Black	-0.01	(0.031)	<0.01	(0.005)	-0.04*	(0.019)	<0.01*	(0.002)	5348	0.828
	Females	0.01	(0.015)	<-0.01	(0.002)	<-0.01	(0.030)	<0.01	(0.002)	5859	0.829
	Males	-0.01	(0.014)	<0.01	(0.002)	-0.02	(0.025)	<0.01	(0.003)	5710	0.827
English Language Arts Test Scores	Overall	-0.01	(0.015)	<0.01	(0.002)	-0.03	(0.021)	<0.01	(0.002)	11605	0.732
	White	0.08	(0.053)	-0.01+	(0.003)	-0.08*	(0.038)	0.01*	(0.004)	1022	0.735
	Black	-0.03	(0.026)	0.01	(0.004)	-0.02	(0.021)	<0.01	(0.002)	5354	0.729
	Females	-0.01	(0.016)	<0.01	(0.002)	-0.05+	(0.030)	<0.01+	(0.002)	5881	0.728
	Males	<0.01	(0.019)	<0.01	(0.003)	-0.01	(0.028)	<0.01	(0.003)	5716	0.719
Ever Suspended	Overall	0.01	(0.007)	<-0.01	(0.001)	-0.02*	(0.011)	<0.01*	(0.001)	11758	0.362
	White	<0.01	(0.041)	<-0.01	(0.003)	-0.01	(0.019)	<-0.01	(0.002)	1029	0.353
	Black	0.01	(0.015)	<0.01	(0.002)	-0.01	(0.017)	<0.01	(0.001)	5438	0.366
	Females	0.01	(0.008)	-0.00+	(0.001)	-0.01	(0.012)	<0.01	(0.001)	5947	0.326
	Males	0.01	(0.010)	<0.01	(0.001)	-0.03+	(0.016)	<0.01+	(0.002)	5801	0.354
Log Number of Days Absent	Overall	-0.02	(0.014)	<0.01	(0.002)	-0.03+	(0.017)	<0.01+	(0.002)	11758	0.624
	White	0.02	(0.047)	<-0.01	(0.004)	<-0.01	(0.038)	<0.01	(0.004)	1029	0.621
	Black	<0.01	(0.018)	<-0.01	(0.003)	-0.02	(0.021)	<0.01	(0.002)	5438	0.609
	Females	-0.02	(0.019)	<0.01	(0.003)	-0.04	(0.030)	<0.01	(0.002)	5947	0.630
	Males	-0.02	(0.015)	<0.01	(0.002)	-0.03	(0.021)	<0.01	(0.003)	5801	0.619

Notes: Standard errors are reported in parentheses and clustered at the school-grade level. Samples are restricted to students with outcome data. All models also include student, year, and grade fixed effects. We standardize the test scores across all students in Boston Public Schools.
+ p<0.10,*p<0.05,**p<0.01