



A Classroom Observer Like Me: The Effect of Demographic Congruence Between Teachers and Raters on Observation Scores

Olivia L. Chi

POLICY ISSUE

State education agencies and school districts across the country are prioritizing the goal of diversifying the teacher workforce. Efforts to recruit, hire, develop, and retain underrepresented teachers are growing alongside evidence that documents both short- and long-term benefits to students who are assigned to demographically similar teachers. Administrators can play an important role in efforts to close the diversity gap. They commonly act as key decision-makers in the processes of hiring, developing, evaluating, and retaining teachers, and they frequently rely on their subjective judgments to make high-stakes human resources decisions.

To further understand the challenges of diversifying the teacher pipeline, I examine how race and gender dynamics influence administrators' subjective assessments of teachers in the context of classroom observations. Specifically, I ask whether teachers receive higher classroom observation scores as a result of sharing race or gender with their observers, who are typically school-based administrators. On a broad scale, if teachers benefit from being evaluated by observers who share their demographic background, this could place non-White and male teachers at a disadvantage, since both groups are underrepresented in the education sector in the United States. In line with previous research, I find that race and gender matches between teachers and their administrators/observers confer a small but meaningful advantage in teachers' observation scores. These results raise fairness concerns for teachers whose demographics are not reflected by any of their observers.

STUDY DATA AND FINDINGS

I use data from a large school district in North Carolina spanning the school years 2013-2014 through 2017-2018. The data includes information from the North Carolina Teacher Evaluation Process (NCTEP), as well as demographic

KEY FINDINGS

Classroom Observer

Teachers, on average, experience small positive increases in their scores from sharing race or gender with their classroom observers.

^ 8-10% of the average gain in **observation scores** demonstrated by novice teachers after an additional year of teaching

data for both teachers and evaluators. During this time period, teachers were observed two to three times per year through the NCTEP as required by the state. Nearly all observations were conducted by administrators at the school site. The analytic sample is restricted to teachers and administrators who identify as either Black or White, as these are the only racial groups for which I have sufficient data to draw meaningful conclusions.

I find that teachers, on average, experience small positive increases in their scores from sharing race or gender with their classroom observers. The magnitudes of these increases are roughly equivalent to 8%-10% of the average gain in observation scores demonstrated by novice teachers after an additional year of teaching. I also explore whether these benefits can be explained by shared experiences between observers and teachers, such as similar prior teaching assignments, attendance at the same university, or a history of working together at the same

school. I do not find evidence that the race- or gender-match effects are explained by these commonalities.

POLICY IMPLICATIONS

Though this study is unable to uncover the underlying mechanisms for the effects of race- and gender-match on classroom observation scores, the results speak to a growing body of research that suggests that observation scores can be influenced by factors that are beyond teachers' control. District personnel may want to consider whether a teacher benefited from such demographic matching when interpreting the scores from classroom observations. In addition, those designing teacher evaluation systems should explore and investigate best practices for how leaders can account for such factors when using observation scores in the decision-making process.

More broadly, the results prompt further investigation into how race and gender influence the subjective assessments of teachers throughout the

teacher pipeline. The dynamics that operate in the context of classroom observations may also play a role in other stages where administrators rely on subjective assessments of teachers to make human resources decisions. Future research ought to examine how administrator and teacher demographics impact the assessments of teachers during the processes of hiring applicants, renewing contracts, and making recommendations for promotions to leadership positions. The presence of similar race and gender dynamics may disadvantage underrepresented teachers and applicants, who are less likely to be assessed by demographically similar administrators. Uncovering whether and in what contexts these dynamics play a role can highlight focus points for initiatives aiming to diversify the teacher pipeline.

FULL REPORT

For the complete working paper, visit wheelockpolicycenter.org.

OUR MISSION

The Wheelock Educational Policy Center (WEPC) conducts and disseminates rigorous, policy-relevant education research in partnership with local, state, and federal policymakers and stakeholders to improve educational opportunities and holistic outcomes for underserved students.

www.wheelockpolicycenter.org
wheelockpolicy@bu.edu



Boston University Wheelock College of Education & Human Development
Wheelock Educational Policy Center

