Teacher bias: A discussion with special emphasis on gender and STEM learning

Ian Thacker

University of Texas at San Antonio

Yasemin Copur-Gencturk

University of Southern California

Joseph R. Cimpian

New York University

Book chapter accepted for publication in T. L. Good & M. McCaslin (Eds.), *The Routledge Encyclopedia of Education: Educational Psychology Edition.* Routledge.

An important feature of human reasoning is to make quick, knee-jerk decisions based on new information (Kahneman, 2011). However, some of this reflexive decision-making is shaped by social context and can ultimately serve as implicit biases that favor privileged social groups and/or harm marginalized social groups. Biases are ubiquitous in society and arise in many contexts, including in schools and the workplace. Even well-intentioned individuals who consciously avoid discriminatory behaviors can enact unconscious biases in their daily lives, which in turn can have serious consequences for marginalized social groups, and teachers are not impervious to such biases. The focus of this chapter is on the sources of and influences of teachers' explicit and implicit stereotypical perceptions, and their implications for STEM teaching and learning. We would like to note early on that our aim is not to blame teachers for larger social problems, but rather, to call attention to pervasive stereotypes that exist in our society that can potentially shape well-intentioned teachers' perceptions and reactions to their students.

Teachers' perceptions comprise one of many factors that can influence students' self-perceptions, beliefs, and achievement. Despite competing hypotheses from economics, social psychology, neuroscience, and sociology regarding different factors that might drive race and gender disparities in educational outcomes (such as arguments emphasizing biological vs sociocultural causes of gender disparities; Halpern et al., 2007), evidence from multiple disciplines is converging on the idea that teachers' conscious and unconscious bias remain important factors contributing to inequitable student achievement outcomes, particularly in mathematics-intensive STEM fields such as computer science or engineering (for reviews of how teacher race- and gender-bias comprises one of many contributors disparities, see Ceci et al., 2009; Cheryan et al., 2017; Dixson & Rousseau, 2005; Wang & Degol, 2017, Warikoo et al., 2016).

This chapter provides interested researchers, practitioners, and policy makers with an overview of research on teacher bias and provide direction for those intending to create interventions to mitigate teacher bias and close these educational gaps. Namely, we summarize research on teacher bias including its history, current empirical research identifying teacher bias, the impacts on students, potential moderators, and implications for teacher education. Following this broad overview, we also include a section providing an in-depth illustration of how bias literature can help us understand a prominent issue: teacher gender-biases as they relate opportunities for women in STEM. Also note that, while the focus of this chapter is on teachers' implicit and explicit biases, we also draw evidence from sources pertaining to other populations as well.

Background and History

Understanding the sources of teacher bias and how it might impact students' performance, self-perceptions, and educational trajectory is important for creating effective interventions to mitigate teacher bias. Early research from economics, social psychology, neuroscience, and sociology begin to indicate the many possible sources and forms of bias among adults and how it might negatively harm marginalized populations exposed to this bias. A recurring theme throughout is a distinction between explicit, controlled forms of bias and implicit, automatic bias.

Economics Perspectives

Decades of research in microeconomics considers how individuals make decisions—who to hire, where to live, and how to rank students in a classroom. Such research sometimes views biases in terms of "taste-based," statistical, and attention-based discrimination (for a review, see

Bertrand & Duflo, 2016). Theory of *taste-based discrimination* posits that discrimination is perpetuated by an individual's explicit prejudices and dislikes for social groups and organizations and manifest as discriminatory hiring decisions, even if those decisions come at a monetary cost (Becker, 1957). Similarly, teachers sometimes (though rarely) explicitly express negative attitudes about certain social groups (see Explicit Bias section for evidence of this claim).

In contrast, *statistical discrimination* is a more subtle form of discrimination that arises in ambiguous situations in which there is a lack of information and a decision maker falls back on social information, such as the average ability of a racial or gender group, to signal unknown information (e.g., Arrow, 1973; Phelps, 1972). For example, a teacher who is asked to judge the ability of a student based on an example of the student's work may feel uncertain about the student's ability if the work is partially correct but partially incorrect. Given such ambiguous information about a student, a teacher may rely on the race or gender of that student to signal unknown information about the student's ability.

Another emerging hypothesis in the economics literature is that of *attention-based discrimination* (Bartos et al. 2016) positing that demographic information triggers different levels of attention within a decision maker. For example, when grading student work, teachers may devote more attention to understanding and providing feedback on male students' work compared with female students' work.

Social-Psychology Perspectives

Early research on *Self-Fulfilling Prophecy* explored how erroneous expectations of students held by teachers can be subtly communicated to and internalized by students in their classroom, leading them to perform at levels consistent with those expectations (see e.g., Good et al., 2018 for a review). As one of the earliest observational studies on teacher expectations, Brophy and Good (1970) used a dyadic coding system of teacher-student interactions to show that first-grade teachers interacted differently with students who they had identified as more and less academically capable. Namely, in three of the four classrooms analyzed, teachers engaged with students in ways that were consistent with their expectations in that they demanded higher performance from students they had identified as having higher capabilities, asked more questions, praised them more for successes, seated students by ability level, and were more likely to accept poor performance from those they held low expectations for. The authors also found that teachers generally had more interactions with boys, direct more evaluative comments to boys, and criticized boys more frequently than girls, particularly those who were perceived as having low ability. Brophy and Good (1974) later replicated their findings in an analysis of nine classrooms, additionally showing that some teachers acted on their expectations of students more than others (over-reactive vs proactive), and also reviewed literature supporting the idea that student characteristics (e.g., race, gender, student handwriting, and other ways that students present themselves to teachers) may be linked to such reactiveness. In the five decades following these studies, research on teachers' expectations have been extended to investigate whether gender, race, and class might serve as signals that trigger different teacher expectations—generally finding that some teachers communicate different expectations to female students through verbal and non-verbal behavior (see Good et al., 2018 for a review). This early research on teacher expectations set the stage for related work on specific instructor behaviors and environmental factors that might trigger of self-doubt among stereotyped groups of students.

Research on *Stereotype Threat* (Steele, 1997) describes a process in which teacher biases might shape disparities in STEM classrooms. Stereotype threat is the phenomenon in which

subtle instructor behaviors or information presented in learning materials trigger feelings of self-doubt, stress, and anxiety among stereotyped learners—hindering their performance on achievement tests. In a series of studies, Steele (1997) presented the effects of triggering negative stereotypes that pertain to undergraduate women and students of color while taking performance exams. Undergraduate women who were strong in math were found to perform worse than men on the mathematics portion of a standardized math exam when told that the test produced gender differences (Spencer et al., 1999), and similarly students of color underperformed their normal ability when told that the exam was a measure of intelligence (Steele & Aronson, 1995). These findings highlight the ways that educators and learning materials can unintentionally trigger feelings of self-doubt among their students.

Another body of research explores how stereotypes and attitudes towards specific social groups do not always reflect conscious thought and sometimes manifest as automatic assumptions about stereotyped social groups (e.g., Devine, 1989; Greenwald & Banaji, 1995). These implicit biases have been measured in a number of laboratory experiments showing that people sometimes enact biases through automatic associations that reflect cultural stereotypes and negative attitudes towards social groups (e.g., Greenwald, McGhee, & Schwartz, 1998; Nosek, Hawkins, & Frazier, 2011). The *Implicit Associations Test* (IAT; Greenwald et al., 1998) is a particularly popular means of capturing implicit stereotypes and attitudes through automatic associations. The IAT captures implicit attitudes by identifying whether individuals tend to more quickly associate positive words (e.g., "good") with words or images signaling a given social group (e.g., images of White faces) and associate negative words (e.g., "bad") with images of a comparison group (e.g., images of African American faces). Similarly, the IAT can capture implicit stereotypes to show that many individuals are, for example, quicker to associate boys and science and math compared with girls. Such implicit stereotypes are thought to reflect people's assumptions about academic abilities of minoritized groups and can manifest as unconscious patterns of thought to influence people's decisions, also known as implicit cognition (Greenwald & Banaji, 1995). The IAT is one of the more popular methods to capture implicit bias, though many other techniques have been used to varying degrees of success (for a review of experimental measures, see Nosek et al., 2011). We should note, however, that some argue that implicit measures such as the IAT may reflect cultural stereotypes embedded in one's surroundings rather than biases of the individual (Arkes & Tetlock, 2004; Payne et al., 2005) as reflected in low test-retest reliability among individuals on some implicit measures of bias (Payne et al., 2017). Namely, implicit bias scores appear to be robust and consistent on aggregate levels (e.g., countries and states), and within individuals over decades, but relatively inconsistent within individuals from week to week, potentially indicating that implicit bias may be an indicator of biases and associations that are situated in context and widespread in society at large rather than something indicative of the individual who demonstrates bias (Payne et al., 2017).

Neuroscience Perspectives

Research in Neuroscience has further improved our understanding of explicit and implicit bias. For example, evidence suggests that a distinct region of the brain regulates automatic, unconscious, "fast" responses (the amygdala) which is associated with implicit bias while another region (the frontal lobe) regulates intentional, conscious, "slow" responses (for reviews, see Amodio, 2014; Kahneman, 2011). These mental processes sometimes result in biases as our brains naturally form relationships in our observed world and construct associations between concepts, such as positive or negative associations between social groups or characteristics. Another important contribution of neuroscience research is that automatic "fast" responses

associated with implicit bias are learned fear responses that are grounded in and reinforced by lived experience. That is, implicit bias is not inborn. People consciously and unconsciously store information from their experiences in their brains which later influences both automatic and intentional decisions.

Sociology Perspectives

Sociological research on organizational factors has contributed to our understanding of how teachers' or peers' biases might affect students (for reviews, see Pager & Shepherd, 2008; Petersen & Saporta, 2004). For example, implicit bias may contribute to processes of forming social networks and groups that share a common identity or interest ("ingroups") in educational settings. Individuals that are not among the ingroup ("outgroups") are then left with fewer networking opportunities which are linked with important outcomes, such as hiring and promotion decisions (Loury, 2009). For instance, teachers who subtly communicate that boys have innate mathematical ability may inadvertently reinforce masculine ingroup distinctions about who belongs in STEM and who does not, potentially alienating girls from participation in social networking opportunities, such as after-school study groups or math clubs.

Further, researchers in sociology have indicated that structural factors such as social norms and policies can shape and reinforce biases and direct their impacts on people's choices and behaviors (Pager & Shepherd, 2008). This body of work argues that existing inequality in educational outcomes and sorting in the workforce by gender and race is a product of preexisting implicit and explicit discrimination and is perpetuated and legitimized as individuals are exposed to these biases and consequences of the biases (e.g., Loury, 2003). For example, teachers' biased messages about mathematical ability can prime students to interpret their own failures in mathematics as being a natural consequence of their membership to a disadvantaged social group. That is, exposure to bias can disadvantage outgroups and trigger self-fulfilling prophecies leading them to conform to stereotypical and biased expectations (Loury, 2003).

Current Research on Teacher Bias

Decades of research across multiple disciplines support the idea that biases manifest either explicitly as intentional, conscious cognitions and behaviors or implicitly as automatic, unconscious evaluations and associations (e.g., Greenwald & Krieger, 2006; Kahneman, 2011). Implicit and explicit biases are widespread, held by a large proportion of individuals across the globe (Coutts, 2020; Miller, Eagly, & Linn, 2015, Nosek et al., 2009)—including some teachers. In this section we present current empirical evidence of the presence and impacts of these forms of bias among teachers; and highlight important measures of teacher bias, the impacts of teacher bias on student outcomes, and interventions intended to mitigate teacher bias.

Explicit Bias

Explicit biases refer to intentional, conscious, discriminatory attitudes and behaviors that are under control of individuals and require controlled and effortful processes and are often captured by simply asking people about their attitudes and beliefs about various social groups (e.g., Greenwald & Krieger, 2006). In other words, explicit biases are social feelings that people are able to articulate and are willing to share. Explicit bias can also vary in terms of the object of discrimination (e.g., attitudes toward multicultural classrooms versus attitudes toward female students) and can reflect attitudes (positive or negative evaluations) or stereotypes (e.g., gender- or race-specific beliefs about a content domain).

Measurement of Explicit Bias

Assessment of teachers' explicit bias relies heavily on the use of questionnaires and self-report data. For example, self-reported surveys of sexism and racism are often used with adults

to capture explicit bias and ask individuals to rate their agreement from statements that range from overt bias (e.g., "Women are generally not as smart as men") to more covert statements that indicate prejudice, such as denial of continued sexism and racism in society with use of the Modern or Symbolic Sexism and Racism scales (e.g., "discrimination against black people is no longer a problem in the United States"; Henry & Sears, 2002; McConahay, 1983; Swim et al., 1995). Another, less often used means of capturing explicit biases is called Q sorting in which participants rank a list of statements in terms of the extent to which they are "like me" or "unlike me" (Stephenson, 1935; Yang & Montgomery, 2013). However, one issue with measuring explicit bias is that, by definition, explicit bias is intentional and under the control of the individual and may therefore be easily adjusted to reflect social norms rather than one's "true" biases.

Evidence on Explicit Teacher Bias

Generally, research on teachers' explicit biases reveal mixed findings, though they tend to suggest that teachers generally report low levels of explicit bias across different contexts. When concerning teachers' attitudes toward racial and ethnic groups, evidence reveals that about 33.3% of US teachers reported pro-White racial attitudes ("warmer" feelings towards White people when compared with Black people), attitudes that were negligibly more positive than those reported by the general public (d = .0002; Starck et al., 2020). Quinn (2017) showed that, although some educators hold negative racial attitudes towards people of color, they tend to give more politically liberal explanations for disparities, hold fewer negative stereotypes, and that such explicit beliefs have decreased over time. When concerning teachers' beliefs about students, survey results suggest that, elementary and middle school math teachers in the USA mostly report low levels of agreement with gender-specific stereotypes about mathematics (Copur-Gencturk, Thacker et al., 2020; Nurnberger et al., 2016; Carlana, 2019).

Teachers' explicit race and gender biases have been shown to be associated with other sets of beliefs and intentions. For example, although teachers tend to overwhelmingly disagree that "boys tend to be smarter than girls at math," the minority of teachers who agree with such statements also hold essentialist beliefs that social categories are natural entities (Nurnberger, et al., 2016) and believe that innate ability or "brilliance" is required for success in mathematics (Copur-Gencturk, Thacker, et al, 2020). Explicit attitudes have also been found to predict behavioral intentions (Levins et al., 2005) though not necessarily behaviors themselves.

However, despite evident interrelations between teachers' beliefs and their behavioral intentions, findings linking explicit bias to teachers' actual behaviors, student outcomes, and even other forms of bias (e.g., IAT scores) yield very few statistically significant relationships (e.g., Carlana, 2019; Glock et al., 2013; Nurnberger, et al., 2016). Explanations for the disconnect between explicit bias and other measurable outcomes vary, one explanation being that there is social desirability bias in self-reported biases, another being that explicit bias is essentially a different construct than implicit forms of bias (Greenwald et al., 2009; Perugini et al., 2010), which are substantially more predictive of teachers' judgments, behaviors, and student outcomes.

Implicit Bias

In contrast with explicit bias, *implicit biases* are the unconscious, automatic associations and attributions held toward social groups that individuals aren't necessarily aware of having (e.g., Blair et al., 2015; Greenwald et al., 1998). A person may truthfully express that they do not hold prejudiced beliefs, yet may still hold implicit biases that steer their judgments and behaviors. For example, a teacher who is assessing student work and says that they "don't see

race" may still unintentionally devote more attention and time to assessing the work of a particular race or ethnicity.

Compared with explicit attitudes, implicit attitudes must be measured indirectly, indicators of which have been shown to predict teacher judgments as well as actual student achievement. We review experimental and observational evidence captured by four indirect measures of teacher bias and their impacts on students: Implicit association tests, audit studies, grading bias, and attributional biases.

Implicit Association Studies

Given that implicit bias is covert and unconscious, it must be measured indirectly. As previously noted, the IAT presents users with different words or images that must be sorted into two categories as quickly as possible. The categories represent social constructs (e.g., male-female) and the sorted words represent stereotypes (e.g., "logical") or attitudes ("good" vs. "bad"). For example, individuals who are quicker to associate boys with science and girls with humanities, but slower to associate girls with science and boys with humanities are thought to hold greater implicit "boy-science" stereotypes.

When aggregated to higher levels, such as the country-level, "boy-science" biases among the general public captured by the IAT have been shown to predict gender-based math and science achievement gaps in 8th grade (Nosek et al., 2009), suggesting that widespread implicit biases present in students' environment are linked with the performance of women in STEM. Additional large-scale studies in the USA suggest that teachers hold implicit racial biases at only slightly lower levels than the general public, as indicated by "black-bad, white-good" associations captured by the IAT (Starck et al., 2020). Implicit teacher bias has also been detected in other countries; "boy-science" biases captured using a technique similar to the IAT have been shown to be associated with gender-stereotypical tracking decisions made by German preservice mathematics teachers (Nurnberger et al., 2016), and in Italy, self-selected tracking decisions among students assigned to their classrooms, as well as longer-term impacts on gender-gaps in mathematics achievement (though not in literature; Carlana, 2019).

Audit Studies

Field-based experimental methods using resume audit techniques are a popular means to measure bias in less controlled settings and involve situated decision-making processes. These methods investigate the impact of race or gender signals on hiring decisions by submitting fictitious resumes to real job openings. For example, Bertrand and Mullainathan (2004) found that when different first names were randomly assigned to otherwise identical resumes and sent to potential employers in Chicago and Boston, White sounding names (e.g., "Emily Walsh") were about 50% more likely to receive a call back for an interview compared with a Black sounding name (e.g., "Lakisha Washington"). Similar audit studies have documented race-based and gender-based discrimination in a variety of contexts (for a review, see Bertrand & Duflo, 2016), including a small number conducted with educators. For example, at the higher education level, Milkman, Akinola, and Chugh (2012; 2015) found that emails sent to STEM professors from fictitious prospective doctoral applications were more likely to get a response, and received faster responses, when coming from a White male student as compared emails sent from female and nonwhite students. Similarly, Moss-Racusin, Dovidio, Brescoll, Graham, and Handelsman (2012) presented science faculty with fictitious lab assistant applications, finding that male applicants were rated as more capable than female applicants among both male and female faculty.

At the K-12 level, our team (Copur-Gencturk, Cimpian, et al., 2020) conducted an experimental study inspired by the audit method. Namely, we investigated biases of elementary and middle-school mathematics teachers when evaluating fictitious student math work that were randomly assigned first names that varied by race and gender. We found that, even though the names of the students did not predict teachers' evaluations of the *correctness* of the student work, significant race and gender differences were found in teachers' evaluations of students' mathematical *ability*, with the largest differences being those favoring White male students when compared with Black female students. Biases were particularly evident for partially correct and incorrect student solutions, highlighting again that implicit biases tend to emerge in ambiguous decision-making situations (e.g., Arrow, 1973; Phelps, 1972). Another central finding of this study was that these biases were not exclusively driven by the White teachers. Specifically, gender biases favoring boy students were found mostly among White female teachers—consistent with our previous work using two nationally representative datasets (Cimpian et al., 2016; Robinson-Cimpian et al., 2014)—and racial biases favoring White students were found mostly among teachers of color.

Grading Bias

Another means of studying teachers' bias in the classroom is by investigating teachers' biased grading patterns of their own students. For example, Lavy (2008) found that teachers in Israel were more likely to give higher exam grades to females when those exams presented names when compared with "blind" exams—though it should be noted that similar grading biases that were captured in the USA appear to largely be explained by non-cognitive skills such as academic engagement (e.g., Cornwell et al., 2013). Grading biases, however, can have long term consequences for students, as Lavy and Sand (2015) found that among 6th grade students, teachers' gender-based grading biases predicted these students' mathematics achievement in elementary school and course enrollment in advanced mathematics in middle school and high school.

Attributional Gender-Bias

A related strand of observational research studies show that teachers in the USA sometimes explain student performance outcomes differently based on social characteristics of the student. Namely, teachers tend to attribute their female students' successes in mathematics to effort but their failures to ability, while with their male students, they attribute their successes to ability and failures to effort (Espinoza et al., 2014; Fennema et al., 1990; Tiedemann, 2000; 2002). One limitation of much of the attributional bias literature, as well as the literature on grading bias, is that there is no way of knowing whether these biases grounded in teachers' own students are based on actual student differences or teachers' implicit or explicit biases.

To address this issue, our team recently explored mathematics teachers' attributional gender bias using experimental methods (Copur-Gencturk et al., 2021), findings of which were largely consistent with the observational research. We asked K-8 teachers to evaluate student mathematics work before randomly assigning them to student performance conditions. For example, some teachers were told that boys had outperformed girls on the assessment while other teachers were told that girls had outperformed boys. We found that teachers attributed gender disparities to differences in effort and opportunity gaps when they were told that girls outperformed boys compared with when boys outperformed girls, and that teachers with low math anxiety were more likely to attribute ability to gender differences when told that girls outperformed boys. Generally, these experimental findings are consistent with observational evidence showing that teachers tend to attribute girls' (relative) successes to effort when

compared with boys, and also finds that mathematics anxiety may be an important moderator of ability attributions.

Potential Moderators of Teachers' Implicit Bias

Though teachers' implicit bias and their impacts have been identified in a number of contexts, less research has investigated important mechanisms and moderators of these biases. This section showcases research on some of these moderators.

Implicit Theories of Intelligence

Beliefs about whether intelligence is malleable or fixed (also called growth- vs fixed-mindset; or implicit theories of intelligence) are considered to be important factors that predict attributions they make for the successes and failures of others (Boaler, 2013; Weiner, 2005; Yeager & Dweck, 2012). For example, a teacher who believes that mathematical intelligence is fixed and unchanging may attribute a student's failure on an exam to their lack of mathematical ability, infer that efforts to change their skills are futile because this ability is innate, and expend less effort to support this student as a result. Such fixed mathematical mindsets held by teachers tend to be disproportionately applied to girls when they fail and boys when they succeed (Espinoza et al., 2014; Fennema et al., 1990; Tiedemann, 2000, 2002). In a series of studies, college-level instructors holding fixed mindsets about math ability reported being more likely to use "comforting" language—such as "explain[ing] that not everyone has math talent"—to students who were said to be struggling with math, rather than provide them helpful strategies; in turn, this "comforting" language demotivated students (Rattan, Good, & Dweck, 2012). Such instructor feedback signaling fixed mindsets about mathematics can lead to decrements particularly for women's performance and sense of acceptance and belonging in college-level mathematics (Good, Rattan, & Dweck, 2012). In other research, in-service elementary and middle school math teachers' general belief that mathematical ability is innate was found to be associated with gender-specific beliefs that boys but not girls hold such innate mathematical ability (Copur-Gencturk, Thacker, et al., 2020). As such, implicit theories of intelligence may be an important moderator of implicit and explicit bias, particularly for gender-specific ability beliefs.

Mathematics Anxiety

Anxiety may also be an important moderator of teacher bias, particularly in the domain of mathematics. Mathematics anxiety refers to feelings of fear and anxiety that people experience when doing mathematics or considering the prospect of doing mathematics (e.g., Ramirez et al., 2018) and was found in a meta-analysis to be more common among women than men (Hembree, 1990). Research with K-12 mathematics teachers indicates that higher levels of math anxiety among White mathematics teachers was predictive of gender-biased attributions for differential successes of boys and girls (Copur-Gencturk et al., 2021). Additional research suggests that anxiety levels of early elementary female teachers were associated with their female students' beliefs in the stereotype that boys are good at math, which in turn predicted their achievement in mathematics (Beilock et al., 2010). Research on teachers suggests that early elementary female teachers' anxiety levels predicted their female students' beliefs that boys are good at math, which in turn predicted their mathematics achievement gains (Beilock et al., 2010), suggesting that teachers' math anxiety may play a role in the transmission of stereotypes from teacher to student.

A Focus on Mathematics Instructor Bias and Gender-Gaps in STEM

Given the broad overview on the topic of teachers' bias in the classroom, we now take a particular focus on teachers' gender-specific implicit and explicit bias regarding students' potential to be successful in mathematics and how they can help us understand an important

issue: gender-gaps in STEM outcomes. We focus on teachers' potential implicit gender-related biases toward students' abilities to further illustrate that, despite nearly equivalent mean mathematics achievement of boys and girls throughout K-12 education, women's career choices seem to be aligned with existing cultural stereotypes. Indeed, women are underrepresented in STEM fields where innate mathematical ability is considered a prerequisite (Leslie et al., 2015). Furthermore, recent analyses have shown that men choose majors in physics, engineering, and computer science at about four times the rate of women, and surprisingly, that women near the 80th percentile of STEM achievement choose these mathematics-intensive majors at the same rate as men at the 1st percentile (Cimpian et al., 2020). While we acknowledge that several factors play a role in STEM career attainment, here we focus on how teachers' implicit and explicit biases regarding gender-related mathematical abilities could be contributing to factors that steer women away from STEM attainment (e.g., expectations for success, confidence, beliefs about ability, secondary course taking).

Evidence and Impacts of Mathematics Educators' Gender Bias

Teachers' explicit and implicit biases can shape teachers' behaviors that have consequences for students. Explicitly held gender stereotypes about STEM fields and negative ability stereotypes comprise two of the important sources of explicit biases that are hypothesized to have impacts on students (Cheryan et al., 2017). While a literature review of early studies (from 1971 to 1996) highlights studies showing that some K-12 teachers hold explicit gender-stereotypes about mathematics being a male domain (Li, 1999), they fortunately do not seem to be as prevalent today (e.g., Copur-Gencturk, Thacker, et al., 2020). Yet, despite low levels of explicit gender stereotypes overall, associations between explicit gender biases and other sets of beliefs and outcomes tell a different story. For example, the small fraction of K-12 teachers who report that they think mathematics requires innate ability also tend to believe that boys, but not girls, have this mathematical ability (Copur-Gencturk, Thacker, et al., 2020). Such discipline-specific beliefs are also prevalent among postsecondary instructors, particularly within mathematics and other STEM fields, and instructors' beliefs that "brilliance" is required for success in their field are associated with women's underrepresentation in those fields (Leslie et al., 2015). Such messages, that mathematics is a fixed trait, when present in the environment also seem to harm women's sense of belonging in undergraduate mathematics and deter them from pursuing mathematics in the future (Good et al., 2012).

Teachers' implicit gender biases have also been detected in a number of mathematics classroom contexts. Observational classroom studies have found that teachers tend to associate mathematical talent with their male students more often than females and attribute their girl students' successes in mathematics to effort and boys' to ability, and girls' failure to lack of ability but boys' to lack of effort (Espinoza et al., 2014; Fennema et al., 1990; Tiedemann 2000; 2002). Such inequitable attributions for math achievement might explain how a teacher might demand more effort from a boy who underperformed yet accept poor performance from a girl who is perceived as having low and fixed ability, leading to inequitable follow-up support. As such, ability attributions can shape teachers' expectations of girls compared with boys (Fennema et al., 1990) which have been shown to impact classroom interactions, ability groupings in the classroom, inter-peer communication, and ranking of student performance (for a review, see Good et al., 2018).

Using nationally representative longitudinal data of U.S. students followed from kindergarteners through fifth grade, our team found that in order to be rated as equal to boys in math, girls as young as in kindergarten needed to be perceived by their teachers as harder

working and more engaged—this is despite external tests showing the boys and girls had equal math performance (Robinson-Cimpian et al., 2014). These patterns persisted each subsequent year of data collection, with each set of new teachers the students encountering in higher grades rating boys as more mathematically able than equally performing and behaving girls. Using combinations of propensity score matching and instrumental variables analyses, we then provided arguably causal evidence that the teachers' underrating of the girls' math abilities contributed substantially to the widening of the gender gap on later math tests. Twelve years later, the U.S. Department of Education conducted a similar data collection on a new cohort of students entering kindergarten, providing us an opportunity to see if the previous patterns would replicate in the new cohort or if more gender equality would prevail. Again, we found that teachers underrated girls' math abilities starting as early as kindergarten, that this persisted throughout early elementary school, and that this underrating of girls likely contributed substantially to the widening of the gender math gap in favor of boys (Cimpian et al., 2016). In both the earlier and later cohorts, this underrating of girls' math ability is done by an almost exclusively female—and predominantly White—set of early elementary educators.

Experimental and laboratory studies also confirm that teachers hold implicit gender-biases that manifest in tracking decisions and student achievement in mathematics. For example, Nürnberger and colleagues (2016) found that German pre-service elementary teachers' implicit associations between male faces and mathematics predicted their decisions to place hypothetical male students in higher level mathematics when compared with female students, while explicit gender biases were not predictive of these tracking decisions. Similarly, a recent study by Carlana (2020) found that 45% of in-service middle-school mathematics and literature teachers in Italy held implicit gender stereotypes associating boys with science, as measured with a Gender-Science IAT, with stronger male-science associations among mathematics teachers. Stronger biases among mathematics teachers predicted significantly larger gender gaps in mathematics achievement on standardized test scores, higher rates of gender-stereotypical tracking recommendations as students moved on to high-school, and lower self-confidence in mathematics among students who were assigned to their classroom—after adjusting for school cohort fixed effects.

Another experimental study was conducted by our team (Copur-Gencturk, Cimpian, et al., 2020) to study implicit biases that arise when mathematics teachers evaluate student work. We recruited three hundred and ninety elementary and middle school teachers from a southern US state that were participating in a professional development to evaluate 18 student solutions. We found that teachers evaluated student mathematics work differently based on the gender and the ethnicity of the first name that was randomly assigned to appear on that work. Student names varied by three ethnicities (Hispanic, White, African American) and gender (male, female) and solutions were either correct, incorrect, or partially correct. When the students' solution was incorrect, teachers tended to infer that students had greater mathematical ability when a male name appeared on the work compared with a female name. This finding was driven mostly by teachers who identified as White and female. This finding was replicated with a national sample of White and Female teachers, in which we also found that such gender-biases were higher amongst teachers who believed that gender-disparities are not prevalent in US society (Copur-Gencturk et al., 2021). Similarly, Avitzour, Cohen, Joel, and Lavy (2020) recently found that Israeli teachers' implicit gender-based grading biases were more prevalent among teachers who believed that they held no biases.

Such unintentional gender biases among mathematics teachers can lead to long-term gender disparities in mathematics. For example, Lavy and Sand (2015) found that gender-biased grading patterns among Israeli elementary teachers were associated with lower achievement for girls and higher achievement for boys that were assigned to their classroom with the effects of a single teacher lasting through middle and high school as evident in gender differences in advanced level math course enrollment.

Summary

In all, this work shows that mathematics teachers hold implicit and explicit gender-biases that are enacted subtly in the classroom by means of differential expectations of students, feedback, and tracking decisions, all of which accumulate over time to impact girls' and young women's self-perceptions in mathematics, potentially dissuading them from pursuing careers in math-intensive STEM disciplines such as physics, engineering, and computer science.

Needed Research

Despite decades of progress in identifying and describing various forms of bias, more work needs to be done to better understand the mechanisms underlying teacher bias, and research must be conducted to explore and design effective means for mitigating bias. While some empirical research has begun to explore potential moderators of bias (such as mathematics anxiety, fixed mindset, and explicit bias) future research might channel such findings into actionable consequences for teacher education and target those day-to-day practices and contexts that will be most beneficial for students. Namely, more research is needed to explore moderators of bias, particularly in discipline- and classroom-specific teaching contexts, as to provide more information about teacher characteristics associated with bias and under what conditions bias arises as to focus the efforts of teacher educators and designers of teacher professional development. Teachers frequently make quick decisions on the basis of ambiguous information in fast-moving and complex settings—teachers have the complex and challenging job of quickly evaluating and providing feedback to students based on their verbalizations and handwritten work, and are given limited time to assign grades, communicate with parents, and make tracking recommendations for students to advanced courses, gifted talented programs, and specialized learning plans—and unfortunately, such situations are prone to be subject to bias. Such situations in which quick decisions must be made by teachers based on ambiguous, unclear information might be explored in future studies, and may inform school administrators' and policy-makers' decisions to create evaluative situations that reduce levels of ambiguity for teachers in such situations.

Furthermore, understanding the mechanisms underlying biases will be critical for designing interventions created to curb them. While rigorous evidence on specific aspects of teacher education programs that might reduce teacher bias is lacking, there are existing programs that have been successful in supporting other populations of adults as they overcome implicit and explicit biases. For example, Devine and her colleagues (2012) fostered reductions in adults' racial bias after 12 weeks by using a 45-minute intervention designed to treat implicit bias like a bad habit, and to break these unintentional habits by creating awareness and concern about bias within oneself and society and by providing specific context-specific strategies for interrupting bias. Carter and colleagues (2020) synthesized such emerging research to make recommendations for school reform, recommending that teacher training should be integrated into schools, but only as part of a broader strategy to promote diversity and inclusion that involves teachers and other adults, and that trainings should focus on promoting awareness and concern about bias and emphasize just a few clear, context-specific strategies with clear

examples for how to manage bias. However, it should be noted that research on the effectiveness of bias and diversity training methods in the classroom context is currently lacking, highlighting the need for more research in this field of study. Future work might explore the effectiveness of existing interventions and center design of new interventions around principles emerging from the research on implicit and explicit bias.

Lastly, studies might examine and compare explicit and implicit biases held by instructors across primary, secondary, and postsecondary education. Such studies might help researchers better understand what kinds of messages students are receiving throughout their academic career, and how those messages might impact students as they potentially shift over time.

References

Amodio, D. M. (2014). The neuroscience of prejudice and stereotyping. Nature Reviews. *Neuroscience, 15*(10), 670–682.

Arrow, K. (1973). The theory of discrimination. In O. Ashenfelter & A. Rees (Eds.), *Discrimination in labor markets* (pp. 3–33). Princeton, NJ: Princeton University Press.

Arkes, H. R., & Tetlock, P. E. (2004). Attribution of implicit prejudice, or "would Jesse Jackson 'fail' the implicit association test?, *Psychological Inquiry*, 15(4), 257–278.

Avitzour, E., Choen, A., Joel, D., & Lavy, V. (2020). *On the Origins of Gender-Biased Behavior: The Role of Explicit and Implicit Stereotypes* (NBER Working Paper w27818). National Bureau of Economic Research. http://www.nber.org/papers/w27818

Bartoš, V., Bauer, M., Chytilová, J., & Matějka, F. (2016). Attention discrimination: Theory and field experiments with monitoring information acquisition. *American Economic Review*, *106*(6), 1437-75.

Becker GS. 1957. *The Economics of Discrimination.* University of Chicago Press.

Bertrand, M., & Duflo, E. (2016). Field experiments on discrimination. In A. V. Banerjee & E. Duflo (Eds.), *Handbook of economic field experiments* (Vol. 1, pp. 309–393). Amsterdam: North Holland Publishing.

Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences*, *107*(5), 1860-1863.

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review*, *94*(4), 991-1013.

Blair, I., Dasgupta, B., & Glaser, J. (2015). Implicit attitudes. In M. Mikulincer & P. R. Shaver (Eds.), *APA handbook of personality and social psychology: Vol. 1. Attitudes and social cognition* (pp. 665–691). American Psychological Association.

Boaler, J. (2013). Ability and mathematics: The mindset revolution that is reshaping education.
Forum, 55, 143–152.

Brophy, J. E., & Good T. L. (1970). Teachers' communication of differential expectation for children's classroom performance: Some behavioral data. *Journal of Educational Psychology*, 61(5), 365-374.

Brophy, J. E., & Good, T. L. (1974). *Teacher-student relationships: Causes and consequences.* Holt, Rinehart & Winston.

Carlana, M. (2019). Implicit stereotypes: Evidence from teachers' gender bias. *The Quarterly Journal of Economics*, *134*(3), 1163-1224.

Carter, E. R., Onyeador, I. N., & Lewis, N. A., Jr. (2020). Developing & delivering effective anti-bias a: Challenges & recommendations. Behavioral Science & Policy, 6(1), 57–70.

Ceci, S. J., Williams, W. M., & Barnett, S. M. (2009). Women's underrepresentation in science: sociocultural and biological considerations. *Psychological Bulletin, 135*(2), 218-261.

Cheryan, S., Ziegler, S. A., Montoya, A. K., & Jiang, L. (2017). Why are some STEM fields more gender balanced than others?. *Psychological Bulletin, 143*(1), 1-35.

Cimpian, J. R., Kim, T. H., & McDermott, Z. T. (2020). Understanding persistent gender gaps in STEM. *Science*, *368*(6497), 1317-1319.

Cimpian, J. R., Lubienski, S. T., Timmer, J. D., Makowski, M. B., & Miller, E. K. (2016). Have gender gaps in math closed? Achievement, teacher perceptions, and learning behaviors across two ECLS-K cohorts. *AERA Open*. https://doi.org/10.1177/2332858416673617

Copur-Gencturk, Y., Cimpian, J. R., Lubienski, S. T., & Thacker, I. (2020). Teachers' bias against the mathematical ability of female, black, and Hispanic Students. *Educational Researcher, 49*(1), 30–43. https://doi.org/10.3102/0013189X19890577

Copur-Gencturk, Y., Thacker, I., & Cimpian, J. R. (April, 2021). *Teachers' implicit gender bias and the moderating effects of modern sexism and anxiety* [Paper presentation]. American Educational Research Association, Orlando, FL.

Copur-Gencturk, Thacker, I, & Y., Quinn, D. (2020). K-8 Mathematics teachers' overall and gender-specific beliefs about mathematical aptitude. *International Journal of Science and Mathematics Education.* Advance online publication.

Cornwell, C., Mustard, D. B., & Van Parys, J. (2013). Noncognitive skills and the gender disparities in test scores and teacher assessments: Evidence from primary school. *Journal of Human Resources*, *48*(1), 236-264.

Coutts, A. (2020). Racial bias around the world (Working Paper). Retrieved from https://osf.io/39svq/

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology, 56*(1), 5-18.

Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. (2012). Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. *Journal of Experimental Social Psychology, 48*(6), 1267-1278.

Dixson, A. D., & Rousseau, C. K. (2005). And we are still not saved: Critical race theory in education ten years later. Race Ethnicity and Education, 8(1), 7–27.

Espinoza, P., Arêas da Luz Fontes, A. B., & Arms-Chavez, C. J. (2014). Attributional gender bias: Teachers' ability and effort explanations for students' math performance. *Social Psychology of Education*, *17*(1), 105–126.

Fennema, E., Peterson, P. L., Carpenter, T. P., & Lubinski, C. A. (1990). Teachers' attributions and beliefs about girls, boys, and mathematics. *Educational Studies in Mathematics, 21*, 55–65.

Glock, S., Kneer, J., & Kovacs, C. (2013). Preservice teachers' implicit attitudes toward students with and without immigration background: A pilot study. *Studies in Educational Evaluation*, *39*(4), 204-210.

Good, C., Rattan, A., & Dweck, C. S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *Journal of Personality and Social Psychology*, 102(4), 700-717.

Good, T. L., Sterzinger, N., & Lavigne. (2018). Expectation effects: Pygmalion and the initial 20 years of research. *Educational Research and Evaluation,* 24(3–5), 99–123.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: attitudes, self-esteem, and stereotypes. *Psychological review*, *102*(1), 4–27.

Greenwald, A. G., & Krieger, L. H. (2006). Implicit bias: Scientific foundations. *California Law Review, 94*(4), 945-967.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology*, *74*(6), 1464–1480.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of personality and social psychology*, *97*(1), 17–41.

Halpern, D. F., Benbow, C. P., Geary, D. C., Gur, R. C., Hyde, J. S., & Gernsbacher, M. A. (2007). The science of sex differences in science and mathematics. *Psychological Science in the Public Interest, 8*(1), 1-51.

Hembree, R. (1990). The nature, effects, and relief of mathematics anxiety. *Journal for Research in Mathematics Education*, 33-46.

Henry, P. J., & Sears, D. O. (2002). The symbolic racism 2000 scale. *Political Psychology*, *23*(2), 253-283.

Kahneman, D. (2011). *Thinking, fast and slow.* Macmillan.

Lavy, V. (2008). Do gender stereotypes reduce girls' or boys' human capital outcomes? Evidence from a natural experiment. *Journal of Public Economics*, *92*(10-11), 2083-2105.

Lavy, V., & Sand, E. (2015). *On the origins of gender human capital gaps: Short and long term consequences of teachers' stereotypical biases* (Working Paper No. w20909). Cambridge, MA: National Bureau of Economic Research.

Levins, T., Bornholt, L., & Lennon, B. (2005). Teachers' experience, attitudes, feelings and behavioural intentions towards children with special educational needs. *Social Psychology of Education, 8*(3), 329–343.

Leslie, S. J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science*, *347*(6219), 262-265.

Li, Q. (1999). Teachers' beliefs and gender differences in mathematics: A review. *Educational Research, 41*(1), 63–76.

Loury, G. C. (2003). Racial stigma: Toward a new paradigm for discrimination theory. *American Economic Review, 93*(2), 334-337

Loury, G. C. (2009). *The Anatomy of Racial Inequality.* Cambridge, MA: Harvard University Press

McConahay, J. B. (1983). Modern racism and modern discrimination: The effects of race, racial attitudes, and context on simulated hiring decisions. Personality and Social Psychology Bulletin, 9(4), 551–558

Milkman, K. L., Akinola, M., & Chugh, D. (2012). Temporal distance and discrimination: An audit study in academia. *Psychological Science*, *23*(7), 710-717.

Milkman, K. L., Akinola, M., & Chugh, D. (2015). What happens before? A field experiment exploring how pay and representation differentially shape bias on the pathway into organizations. *Journal of Applied Psychology*, *100*(6), 1678–1712.

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, *109*(41), 16474-16479.

Nosek, B. A., Hawkins, C. B., & Frazier, R. S. (2011). Implicit social cognition: From measures to mechanisms. *Trends in Cognitive Sciences*, *15*(4), 152-159.

Nosek, B. A., Smyth, F. L., Sriram, N., Lindner, N. M., Devos, T., Ayala, A., ... & Kesebir, S. (2009). National differences in gender–science stereotypes predict national sex differences in science and math achievement. *Proceedings of the National Academy of Sciences*, *106*(26), 10593-10597.

Nürnberger, M., Nerb, J., Schmitz, F., Keller, J., & Sütterlin, S. (2016). Implicit gender stereotypes and essentialist beliefs predict preservice teachers' tracking recommendations. *Journal of Experimental Education, 84*(1), 152–174.

Pager, D., & Shepherd, H. (2008). The sociology of discrimination: Racial discrimination in employment, housing, credit, and consumer markets. *Annual Review of Sociology 34,* 181-209.

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology, 89*, 277– 293.

Payne, B. K., Vuletich, H. A., & Lundberg, K. B. (2017). The bias of crowds: How implicit bias bridges personal and systemic prejudice. *Psychological Inquiry, 28*(4), 233-248.

Petersen, T. & Saporta, I. (2004). The opportunity structure for discrimination. *American Journal of Sociology, 109*(4), 852-901.

Phelps, E. S. (1972). The statistical theory of racism and sexism. American Economic Review, 62(4), 659–661Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8*(3), 220-247.

Perugini, M., Richetin, J., & Zogmaister, C. (2010). Prediction of behavior. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 255–277). The Guilford Press.

Quinn, D. M. (2017). Racial attitudes of preK–12 and postsecond- ary educators: Descriptive evidence from nationally representative data. *Educational Researcher, 46*(7), 397–411.

Ramirez, G., Shaw, S. T., & Maloney, E. A. (2018). Math anxiety: Past research, promising interventions, and a new interpretation framework. *Educational Psychologist*, *53*(3), 145-164.

Rattan, A., Good, C., & Dweck, C. S. (2012). "It's ok—Not everyone can be good at math": Instructors with an entity theory comfort (and demotivate) students. *Journal of Experimental Social Psychology*, *48*(3), 731-737.

Robinson-Cimpian, J. P., Lubienski, S. T., Ganley, C. M., & Copur-Gencturk, Y. (2014). Teachers' perceptions of students' mathematics proficiency may exacerbate early gender gaps in achievement. *Developmental Psychology*, *50*(4), 1262-1281.

Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*(1), 4-28.

Starck, J. G., Riddle, T., Sinclair, S., & Warikoo, N. (2020). Teachers are people too: Examining the racial bias of teachers compared to other American adults. *Educational Researcher*, *49*(4), 273-284.

Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*(6), 613–697.

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.

Stephenson, W. (1935). Correlating persons instead of tests. *Journal of Personality, 4*, 17 24.

Swim, J. K., Aikin, K. J., Hall, W. S., & Hunter, B. A. (1995). Sexism and Racism: Old-Fashioned and Modern Prejudices. *Journal of Personality and Social Psychology*, *68*(2), 199–214.

Tiedemann, J. (2000). Gender-related beliefs of teachers in elementary school mathematics. *Educational Studies in Mathematics, 41*(2), 191–207.

Tiedemann, J. (2002). Teachers' gender stereotypes as determinants of teacher perceptions in elementary school mathematics. *Educational Studies in Mathematics, 50*, 49–62.

Wang, M.-T., & Degol, J. L. (2017). Gender gap in science, technology, engineering, and mathematics (STEM): Current knowledge, implications for practice, policy, and future directions. *Educational Psychology Review*, *29*(1), 119–140.

Warikoo, N., Sinclair, S., Fei, J., & Jacoby-Senghor, D. (2016). Examining racial bias in education: A new approach. Educational Researcher, 45(9), 508-514.

Weiner, B. (2005). Motivation from an attribution perspective and the social psychology of perceived competence. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 73–84). New York: Guilford Press.

Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist*, *47*(4), 302-314.

Yang, Y., & Montgomery, D. (2013). Gaps or bridges in multicultural teacher education: a Q study of attitudes toward student diversity. *Teaching and Teacher Education,* 30 , 27-37.