

# Loss of Institutional Trust Among Racial and Ethnic Minority Adolescents: A Consequence of Procedural Injustice and a Cause of Life-Span Outcomes

David S. Yeager

University of Texas at Austin

Valerie Purdie-Vaughns

Columbia University

Sophia Yang Hooper

University of Texas at Austin

Geoffrey L. Cohen

Stanford University

This research tested a social-developmental process model of trust discernment. From sixth to eighth grade, White and African American students were surveyed twice yearly (ages 11–14; Study 1,  $N = 277$ ). African American students were more aware of racial bias in school disciplinary decisions, and as this awareness grew it predicted a loss of trust in school, leading to a large trust gap in seventh grade. Loss of trust by spring of seventh grade predicted African Americans' subsequent discipline infractions and 4-year college enrollment. Causality was confirmed with a trust-restoring "wise feedback" treatment delivered in spring of seventh grade that improved African Americans' eighth-grade discipline and college outcomes. Correlational findings were replicated with Latino and White students (ages 11–14; Study 2,  $N = 206$ ).

I am marked by old codes, which shielded me in one world and then chained me in the next.—*Ta-Nehisi Coates (2015), an African American father, in a letter to his adolescent son.*

Support for this research was provided by grants from the National Science Foundation/REESE Division (Award 0723909), the Spencer Foundation (Award 200800068), the W. T. Grant Foundation, the Nellie Mae Education Foundation, and National Science Foundation Grant DRL-1109548 to the last author. Writing of the manuscript was supported by a fellowship from the Center for Advanced Study in the Behavioral Sciences (CASBS) and a William T. Grant Scholars award to the David S. Yeager. The authors are grateful to Eric Bettinger for his assistance obtaining National Student Clearinghouse data and to Hae Yeon Lee for her analyses of the School and Staffing Survey. Data, syntax, and results of alternative model specifications are available online at [osf.io/3hpu8](http://osf.io/3hpu8).

Correspondence concerning this article should be addressed to Geoffrey L. Cohen or David S. Yeager. 108 E. Dean Keeton Stop A8000, Austin, TX, 78712-1043. Electronic mail may be sent to [glc@stanford.edu](mailto:glc@stanford.edu) or [dyeager@utexas.edu](mailto:dyeager@utexas.edu).

[Article corrected on March 31, 2017 after first online publication on February 8, 2017: On page 659, "were white" was inserted into the sentence, ". . . 70% of teachers in minority-serving schools were white (i.e. those with 50% or more non-white students. . .". On page 662, two variations of the same sentence were corrected to read, "In these experiments, wise feedback clarified that the critical feedback originated in the critic's positive motivations rather than bias." On page 673, a duplicate appearance of "send the" was removed from the sentence, "Instead, they continually send the message that their students are capable. . ." On page 674, the reference to a Gallup poll was corrected to be "Jones, J. M. (2015, June 19). In U.S., confidence in police lowest in 22 years. Omaha, NE: Gallup. Retrieved from <http://www.gallup.com/poll/183704/confidence-police-lowest-years.aspx>". Additionally, grammatical changes were made to the article.]

A key developmental challenge in contemporary society is to learn how to monitor, assess, and reassess the trustworthiness of institutions and ultimately make judgments about whether compliance with institutional policies is warranted (Purdie-Vaughns, Steele, Davies, Dittmann, & Crosby, 2008; Spencer, 2006; Tyler & Blader, 2003). Adolescents face this challenge with respect to many institutions, such as police, government, healthcare, or businesses. Here, we focus on one setting where trust discernment comes to the fore and may have long-term consequences: racial and ethnic minority adolescents transitioning through middle school.

On the one hand, it can be costly for an adolescent to trust an institution when it systematically takes advantage of one's group (Cohen & Steele, 2002; Gambetta, 1988). Indeed, in the U.S. schools, African American and Latino youth are disproportionately subjected to mistreatment by authorities, in the form of low expectations from teachers (see Harber, 1998), extreme disparities in disciplinary referrals and suspensions for minor misbehavior (Okonofua & Eberhardt, 2015; Skiba, Michael, Nardo, & Peterson, 2002), and even unwarranted stops and arrests by police (see Tyler, Fagan, &

Geller, 2014). Because trust of authorities is negotiated in the context of social interactions (Lewin, Lippitt, & White, 1939), a natural adaptation to such interactions is to withhold trust.

As illustrated by the quotation above, from Tanehisi Coates (2015) to his 14-year-old son, experiences of injustice can affect one in ways that can make it more difficult to thrive in an institution. That is, racial and ethnic minority youth, experiencing and perceiving bias, may generalize from specific interactions to a mental representation of the institution as an abstract entity. Youth may then demonstrate lower compliance with institutional policies, accelerating a self-reinforcing cycle of punishment and loss of trust (Crocker & Major, 1989; Fagan & Tyler, 2005; Gregory & Weinstein, 2008; Levi, Sacks, & Tyler, 2009; Tyler, Goff, & MacCoun, 2015). Exacerbating the potential for mistrust is the racial divide that characterizes most interactions between teachers and their students in the U.S. According to our re-analyses of the School and Staffing Survey (Coopersmith, 2009), 70% of teachers in minority-serving schools were white (i.e. those with 50% or more non-white students; the figure is 96% in predominately white-student-serving institutions). The tendency to attribute bias is greater across racial divides than within them (Crocker & Major, 1989); thus the demographics of most American schools may act as an affordance of mistrust.

Yet a loss of trust can become a liability if some authority figures wish to enact fair policies, treat racial or ethnic minority students with respect, and help them to thrive (Gregory & Weinstein, 2008; Okonofua, Walton, & Eberhardt, 2016). When students have lost trust, they may be deprived of the benefits of engaging with an institution, such as positive relationships, access to resources and opportunities for advancement, and avoidance of punishment. Thus, racial and ethnic minority youth may be twice harmed by institutional injustices: They both receive the lion's share of the initial punishment, and then may be required to psychologically adapt, through a loss of trust, in a way that prevents them from profiting from instruction and relationships.

This research seeks to understand this predicament by applying a developmental and social-psychological lens (also see Okonofua, Paunesku, & Walton, 2016; Okonofua, Walton, et al., 2016). We use data from an 8-year longitudinal study that tracked White and African American students from sixth grade until entry into college across two cohorts (Study 1). In addition, because our theory is that trust is an issue for any group that faces

negative stereotypes and institutional mistreatment, we also examine data from a 1-year, cohort-sequential study of White and Latino middle school students (Study 2).

Our approach integrates insights from developmental (Killen, Mulvey, & Hitti, 2013; McKown, 2013; Olson & Dweck, 2008) and social-psychological recursive process models (Okonofua, Walton, et al., 2016), resulting in the model shown in Figure 1. We document the mental representations (i.e., institutional trust) that sit in the middle between objective realities (i.e., racial and ethnic minority group differences in the perceived justice of school procedures) and developmental outcomes (i.e., group differences in rates of noncompliance with school rules, or in long-term educational outcomes such as college enrollment). We highlight how these mental representations are shaped and then create social reality, propelling effects forward in time.

Finally, we use a small sample, double-blind randomized experiment to test causality. We illustrate how a timely and credible show of respect from authorities to racial and ethnic minority youth can, during a key developmental window, set in motion an alternative process. This more beneficial process can culminate in improved educational behaviors years later.

### *Institutional Trust Among Adults*

Among adults, what causes institutional trust? Research in social psychology (Tyler, 2006; Tyler & Blader, 2003; Tyler et al., 2015), political science (Levi, 1997; Levi et al., 2009), and sociology (Bryk & Schneider, 2002; Goffman, 1963) has converged on the conclusion that people trust an institution more when they perceive that it is *procedurally just*—that is, that it uses fair processes to make consequential decisions—and when they believe that authorities have *personal regard* for individuals served by the institution—that is, when authorities are respectful and have one's best interest at heart. Notably, people do not have to have been the personal recipient of procedural justice or personal regard to gain or lose trust (Tyler et al., 2014). It is enough to experience it vicariously by means of observation.

What has been missing from much previous research, however, has been a developmental perspective. Indeed, an adolescent developmental lens was a notable absence from recent influential reviews of group disparities in crime and policing (President's Task Force on 21st Century Policing, 2015; Tyler et al., 2015). Our primary theoretical contribution is to offer such a lens.

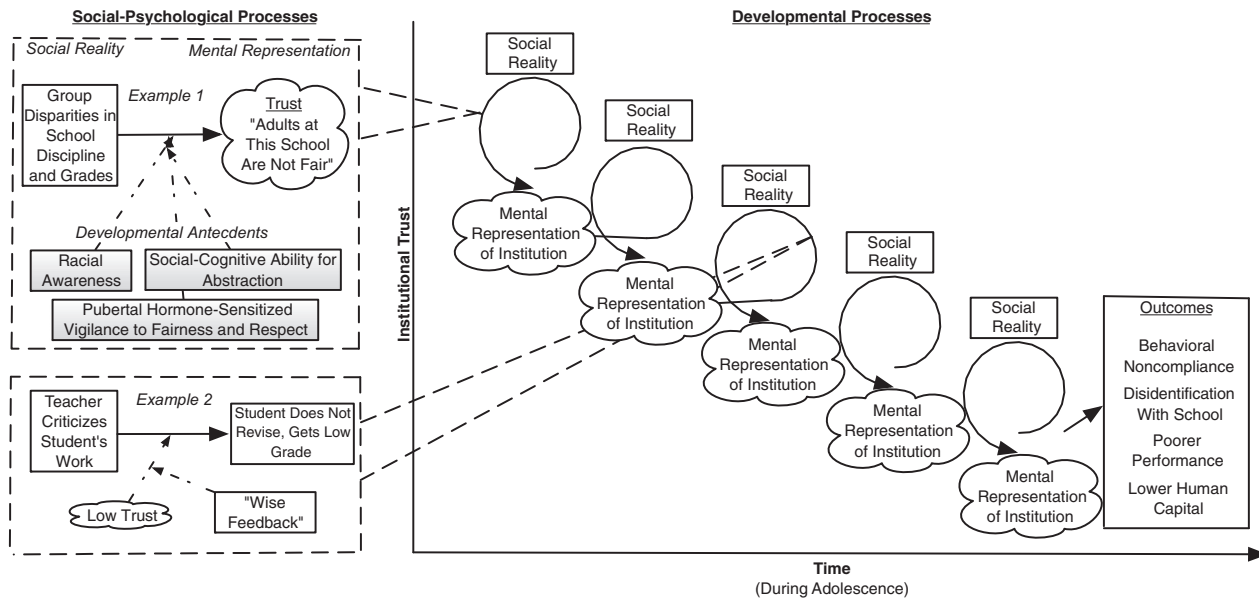


Figure 1. A social-developmental recursive process model of trust discernment during adolescence.

#### *A Developmental Process Model of the Emergence and Consequences of Institutional Trust*

Adolescence has been described as a period of socialization into a belief about the fairness and legitimacy of institutions (Baumrind, 1997; Fagan & Tyler, 2005). Positive and negative encounters with authorities—such as discipline incidents or moments of criticism—are “teachable moments,” for good or ill (Baumrind, 1997; Fagan & Tyler, 2005; Tyler et al., 2014). Of course, even young children are able to discern whether to trust another person (Vanderbilt, Liu, & Heyman, 2011) and are aware of racial and ethnic group membership (see Brown & Bigler, 2005). Why then might adolescence be a special period for institutional trust discernment?

To trust—or distrust—an *institution*, one must notice and assign generalized meaning to authorities’ negative treatment of members of one’s group and attribute it to a bias in the institution as an abstract actor (the school, or school in general) not just to an individual actor (a specific teacher). Several developmental antecedents make adolescents prepared to do this (see the top left panel in Figure 1).

#### *Changes During Adolescence*

Adolescents are increasingly capable of mentally representing unseen, abstract policies in an institution as the cause of unfair treatment, whereas children may be more attuned to the negative attitudes

of specific authority figures acting as individuals (see Brown & Bigler, 2005; Spencer, 2006). Next, adolescence is a period marked by questioning whether to obey adults’ commands (Laupa, Turiel, & Cowan, 1995) or ignoring them altogether (for an example from neuroscience, see Lee, Siegle, Dahl, Hooley, & Silk, 2014). Third, adolescents may undergo pubertal changes that can heighten a sensitivity to the possibility of being disrespected by adult authorities. The pubertal surge in testosterone and other hormones among boys and girls (see Peper & Dahl, 2013) is thought to increase attention to, reactivity to, and memory for experiences of being disrespected or made to feel low status (see Terburg & van Honk, 2013).

#### *Developments Specific to Minority Youth*

As they mature, adolescents undergo these changes regardless of ethnic group membership. However, by the start of middle school, negatively stereotyped ethnic minority adolescents are also more likely than White peers to be *racially and ethnically aware*—that is, to have conscious appraisals about how different racial and ethnic groups are evaluated and treated by the larger society (Ruble et al., 2004; Sellers, Copeland-Linder, Martin, & Lewis, 2006; Spencer, 2006; Swanson, Cunningham, Youngblood, & Spencer, 2009). Long before they get to sixth grade, children and teachers stereotype African American children as aggressive (Sagar & Schofield, 1980), and children take note. By early

adolescence, African American students report occurrences of others perceiving them as a threat (Sellers et al., 2006), and the overwhelming majority are aware of negative stereotypes about their group's intelligence or behavior by the beginning of middle school, even as their White peers are unaware (Bigler, Averhart, & Liben, 2003; McKown & Weinstein, 2003).

Negatively stereotyped racial and ethnic minority adolescents may therefore enter middle school prepared to attribute unfair treatment to group membership rather than group-irrelevant factors and do so more readily than White peers (Crocker & Major, 1989, McKown & Weinstein, 2003; also see Cohen & Garcia, 2005). This preparedness then meets social reality in the form of disparate punishment and remediation in school (Okonofua & Eberhardt, 2015; see Skiba et al., 2002). Even when students themselves are not subjected to disparate treatment, vicarious experiences of group members offer a window into the potential unfairness endemic to the institution (see Bigler & Liben, 2006).

#### *Recursive Process*

Like a stone rolling down a hill that triggers an avalanche, the loss of trust—especially among members of negatively stereotyped minority groups—could accumulate behavioral consequences over time (see Roberts & Caspi, 2003; Spencer, 2006), as illustrated in the right panel of Figure 1. For members of stereotyped groups, disparities in treatment in a given institution can amplify signals of unfairness and disrespect (Crocker & Major, 1989, Purdie-Vaughns et al., 2008), while obscuring signals of fair treatment (Harber, 1998), accelerating loss of trust, and its potential negative consequences (Okonofua, Walton, et al., 2016).

Social-psychological process models emphasize the importance of *attributions* in this feedback loop (Cohen & Steele, 2002; Okonofua, Walton, et al., 2016; see the left panels in Figure 1). If adolescents experience procedural injustice or disrespect, they may come to expect it. If they expect it, they will tend to see it. If they see it, they will tend to expect more (cf. Lord, Ross, & Lepper, 1979). Seeing and expecting injustice and disrespect, negatively stereotyped ethnic minority adolescents may disengage, defy authorities, underperform, and act out. This process may culminate in greater disciplinary incidents and lower academic attainment, even among ethnic minority adolescents who began middle school motivated and capable of

succeeding (Gregory & Weinstein, 2008; Okonofua, Walton, et al., 2016; Spencer, 2006).

Our process model offers an alternative to deficit models of adolescent behavior problems. It does not posit deficiencies in executive function, cultural mismatches between school contexts and ethnic minority families. Nor does it posit a developmentally deterministic path in which adolescents inevitably rebel against adult authority. These may play a role but the causal action in our model is in the interaction between social reality and psychology. A loss of trust and its behavioral consequences are often reasonable adaptations to an environment of disrespect and procedural injustice (Spencer, 2006; Swanson et al., 2009). This adaptation can have long-term consequences when left unaddressed. However, as a mental representation, trust can be improved through timely intervention (see Olson & Dweck, 2008).

#### *Slowing a Loss of Trust Through "Wise" Strategies That Respect Youth*

What might prevent a loss of trust in adolescence? "Wise" strategies can accomplish this. They convey to students that they will be neither treated nor judged in light of a negative stereotype but will instead be respected and treated as a valuable individual (Cohen & Steele, 2002; Goffman, 1963). If such a signal was offered during sensitive periods for trust formation, it might create accumulating positive consequences, by means of a virtuous recursive cycle rather than a vicious one (Yeager & Walton, 2011).

The term "wise" was originally formulated by Goffman (1963) in his analysis of social stigma. To be "wise" is the act of seeing stigmatized individuals in their full humanity, which enables an openness and honesty when one interacts with them.

Ethnographic accounts of exceptional teachers show how wise practitioners build trust with and motivate members of stigmatized groups (e.g., Lepper & Woolverton, 2002; Treisman, 1992; see Gregory & Weinstein, 2008). Such teachers have been described as "warm demanders" (Vasquez, 1988) and "compassionate disciplinarians" (Irvine, 2003), meaning that they are distinguished by their ability to combine high performance standards with high personal regard for students' well-being. One student described a "wise" teacher this way: "When she talk to you with seriousness, she mean [sic] it, but then she also have a smile like 'I'm on your side'" (Gregory & Weinstein, 2008).

In large-scale survey studies, this combination of high standards and personal care predicted higher



achievement for negatively stereotyped and socioeconomically disadvantaged youth (Gregory & Weinstein, 2004; Shouse, 1996). Furthermore, in a daily diary study, African American students who reported experiencing *both* high expectations and feelings of personal care were the most likely to report that they trusted their teacher (Gregory & Weinstein, 2008, Study 2)—regardless of the teacher’s racial and ethnic group.

Research has distilled the practices of wise mentors to investigate them experimentally. Cohen, Steele, and Ross (1999) showed that when African American students were required to endure tough criticism on an essay, they benefitted from learning that the teacher was critical not because of bias but because of his stated desire to hold them to a higher standard (*procedural justice*) and his belief that they were capable of meeting that higher standard (*personal regard*). This message—called “wise feedback”—increased African American college students’ motivation and reduced their attributions of bias (Cohen et al., 1999). In these experiments, wise feedback clarified that the critical feedback originated in the critic’s positive motivations rather than bias. By changing attributions in this manner, the intervention turned what might have otherwise been seen as negative feedback into positive feedback. This wise feedback method was later tested in a behavioral field study that affected African American seventh-grade students’ institutional trust and motivation (Yeager et al., 2014). Yeager et al. (2014) showed that a single hand-written note from students’ social studies teacher, appended to an essay they had written, and pithily but credibly conveying the teacher’s high standards and belief in students’ capacity to reach them, increased from 17% to 72% the proportion of African American adolescents who revised their critiqued essays a week later (covariate-adjusted values). Furthermore, 2.5 months later, the note halted the semester-over-semester decline in trust for African American seventh graders who, at baseline, had begun to mistrust school.

In this research (Study 1), we analyze additional longitudinal data from the same data set originally reported by Yeager et al. (2014; Studies 1 and 2). We look at the implications of an experience of wise feedback on African American adolescents’ subsequent disciplinary infractions and eventual 4-year college enrollment.

#### *This Research*

We test whether disciplinary outcomes in school, and the sense of procedural injustice they give rise

to, predict a loss of institutional trust for negatively stereotyped racial and ethnic minority adolescents (Figure 1, top left panel), and whether this loss in turn predicts greater awareness of procedural justice, starting a feedback loop (Figure 1, right panel). We assume but do not measure developmental antecedents that can facilitate this process, such as racial and ethnic awareness, social-cognitive ability, and the start of puberty (Figure 1, top left panel). We then test whether middle school adolescents’ trust predicts their later behavioral outcomes—their discipline incidents in eighth grade and their eventual enrollment at a 4-year college (bottom right of Figure 1).

Finally, as noted, we take advantage of a randomized experimental intervention embedded in this correlational study, to test causality. Can wise critical feedback from a teacher, which has already been shown to sever the effect of academic mistrust on subsequent willingness to revise an essay and later mistrust (Yeager et al., 2014, Studies 1 and 2), set in motion a virtuous recursive cycle? We explored whether a wise hand-written note from a teacher, known to convey respect and build trust (cf. Gregory & Weinstein, 2008), might set in motion a process that leads to greater behavioral compliance the subsequent year as well as greater eventual enrollment at a 4-year college (also see Okonofua, Paunesku, et al., 2016).

## Study 1

### *Method*

#### *Participants*

*Survey sample.* A total of  $N = 277$  students in two consecutive cohorts (Fall of 2004 and 2005) were recruited from a middle-class to lower-middle-class public middle school in the northeast region of the United States and surveyed twice yearly from sixth to eighth grade. The overall school population was split evenly between African American and White students. Twenty-two percent were eligible for free or reduced-price lunch, an indicator of socioeconomic status. Fifty-two percent of the students were female; 48% were African Americans, and 52% were White. Not all students completed all measures, so degrees of freedom varied across analyses.

*Experimental subsample.* In the spring of seventh grade in each cohort, a subsample of 44 students (88 total) with “B” or “C” grades participated in a randomized experimental intervention in their

social studies classrooms. The focus on “B” and “C” students was motivated by the fact that the intervention conveyed to students that they were able to reach a higher standard. This message, it was expected, would lack credibility in the eyes of very low-performing students and seem redundant in the eyes of very high-performing students, who were already, by and large, reaching a higher standard (Yeager et al., 2014).

All social studies teachers were white. This matches the modal experience for students in the U.S., as noted previously.

In each cohort,  $n = 11$  students from each racial group (White and African American) were assigned either to a treatment condition or placebo control condition, yielding 22 total per racial group per condition. These constituted the “experimental subsample” (for extensive detail on this subsample, see Yeager et al., 2014, Studies 1–2).

### Procedure

*Survey sample.* Students were surveyed in the spring and fall of sixth, seventh, and eighth grades (six total surveys). At the end of the middle school years, data on discipline incidents and school grades were collected from the school registrar. College enrollment data (whether students enrolled in the year after high school and whether it was a 4-year college or not) were obtained 6 years after middle school through the National Student Clearinghouse (NSC).

*Experimental subsample.* The “wise feedback” intervention was designed to be an antidote to mistrust of teachers. No intervention could fully dispel concerns about racial bias. Instead, this treatment was aimed to interrupt the reciprocally reinforcing relationship between perceived bias and mistrust (Figure 1). Because the process depends on a continual feedback loop, even a minor interruption could have benefits that persist and compound with time. The intervention altered students’ construals of an acute experience with the potential to increase mistrust: the receipt of teachers’ critical feedback on an essay they had written. Such feedback could plausibly be attributed to either negative or positive factors, such as bias or a belief in one’s ability to reach a higher standard (Cohen & Steele, 2002). The treatment sought to promote the latter judgment.

In spring of seventh grade, in the context of their social studies class, students wrote an essay about a personal hero. Their initial draft received thorough critical feedback from their teachers, accompanied by whatever encouragement the teacher felt inclined to give. Before students received their

edited essay, the research team appended one of two notes to the essays of the participating students. Both notes had been hand written by the teacher in advance. In a *placebo* control condition, the note stated, “I’m giving you these comments so that you’ll have feedback on your paper,” which fulfills the grammatical expectation to provide an explanation but conveys no information relevant to the trustworthiness of the criticism. In the *wise feedback* condition, the note stated, “I’m giving you these comments because I have very high expectations and I know that you can reach them.” A blank piece of paper covered the notes and edited essays to keep peers and teachers blind to condition.

Over the next week, students had the opportunity to revise their essays, which most did. Yeager et al. (2014) previously reported that the randomly assigned “wise feedback” note increased African American students’ likelihood of revising their essays and improved the quality of their revisions. For instance, it doubled the number of teacher-suggested edits that students implemented. Furthermore, 3 months posttreatment, African American students with low levels of baseline trust showed a less steep decline in trust in the treatment condition than in the control condition.

### Survey Measures

*School trust.* Trust was assessed with six questions (e.g., “I am treated fairly by teachers and other adults at my school,” “Students in my racial group are treated fairly by teachers and other adults at [school name] middle school;” 1 = *very much disagree*, 6 = *very much agree*). Items were averaged at each wave (all  $\alpha$ s > .70), with higher values indicating greater trust.

*Awareness of bias in enforcement of school policies.* Students reported their belief that racial bias intruded in school disciplinary decisions—and thus provided a measure of procedural injustice—using two items written specifically for this study (“If a Black or a White [school name] student is alone in the hallway during class time, which one would a teacher ask for a hall pass?” and “If a Black and a White [school name] student do something wrong, who is more likely to get in trouble for it?” Rated on a 5-point scale; 1 = *almost always the black student instead of the white student*; 3 = *both of them about the same*; 5 = *almost always the white student instead of the black student*). Only 2% of all participants ever gave a response > 3, which would have indicated *pro*-African American bias. Therefore, the measure is best interpreted as perceptions of *anti*-African American

bias. The two items were averaged, with higher values corresponding to a greater expectation of unjust treatment ( $\alpha > .70$ ). We then reverse coded the composite score, with higher values corresponding to a greater expectation of procedural injustice. The modal response was “about the same,” with over 55% of students at each wave providing a rating of 3. Data could thus be analyzed either in their original continuous form or dichotomized. Although we used the former method, substantive results were unchanged with the latter.

*School records.* From school records we obtained the following data for each student: their *gender*, the number of *disciplinary incidents* in each of the three middle school years (range: 0–34), and grade point averages (GPAs) in core classes (math, science, social studies, and English or reading; range: 0 [F] to 4.33 [A+]) for each academic quarter of each of the three school years. As is common in this kind of administrative data, disciplinary incidents were nonnormal due to large numbers of students with 0 incidents and few with very many ( $> 20$ ). To better normalize the distribution, high outliers were recoded to the 98th percentile and then the overall metric was square root transformed. However, substantive results were unchanged when this was not done (i.e., significant results in the article were significant in the raw data as well). In addition, for the first cohort only, the school provided data on the reason for each disciplinary action.

*College enrollment.* College enrollment data were obtained from the NSC. NSC is a nonprofit database that reports on students receiving financial aid to both private and federal loan providers, and as a result it tracks college enrollment and degree attainment for the vast majority of college students in America. Researchers conducting program evaluation in partnership with schools, as in this case, can obtain participating students' college enrollment data. In the Northeast region of the United States, where nearly all participating students attended college, the coverage rate for African American students attending public 4-year institutions was 99.6%, at public 2-year institutions it was 99.2%, and at private nonprofit institutions it was 97% (Dynarski, Hemelt, & Hyman, 2013).

The primary outcome of interest was *on time enrollment at a 4-year college* (1 = enrolled at a 4-year college the year after high school; 0 = did not), which was the appropriate measure for the middle-achieving (B and C average) students who participated in the experiment. Exploratory analyses also focused on whether students enrolled in college at all; roughly 70% of students in this middle-class

school had records the year after high school, indicating college enrollment. Just over two thirds of these students enrolled in a 4-year college.

*Covariates.* Covariates used in all structural equation models were as follows: students' *gender*, *cohort*, and *prior achievement*. The latter was an average of students' state test scores in the first month of sixth grade and their *prior grades*. To ensure that these achievement measures represented baseline values, the correlational analyses, which began with data from the fall of sixth grade, used fifth-grade GPA, whereas analyses of the experimental treatment, implemented in the spring of seventh grade, used GPA from the fall of seventh grade. College enrollment models, which as described below used seventh- and eighth-grade measures, also controlled for prior (sixth grade) trust.

## Results

The software Mplus was used for all structural models. Full information maximum likelihood was used to estimate parameters without discarding or imputing missing data. For models predicting school discipline (a count variable), Poisson regression was used. Additional detail on all statistical models can be found in the Supporting Information, and syntax is posted online (osf.io/3hpu8).

### *The Emergence of the Middle School Racial Trust Gap*

In a growth curve model, including all students from both racial groups, trust decreased from Grades 6 to 8, slope: unstandardized  $b = -0.13$ ,  $p < .001$ , 95% CI  $[-0.17, -0.09]$ . The model fit the data acceptably:  $\chi^2(29) = 83.45$ ,  $p < .001$ ; root mean square error of approximation (RMSEA) = .08, 90% CI  $[0.06, 0.10]$ ; and comparative fit index (CFI) = .93.

However, trust declined faster for African American students, producing a racial trust gap. Figure 2 shows that there was no significant trust gap between African American and White students in sixth grade, but the gap emerged by fall of seventh grade. Student race significantly predicted the slope from sixth to eighth grades in a latent growth curve model, unstandardized  $b = -0.06$ ,  $p = .028$ , 95% CI  $[-0.11, -0.01]$ . This steeper loss of trust for African American students led to significant race-based differences in trust by seventh and eighth grade.

### *Predictors of the Trust Gap*

*Does age alone account for the trust gap?* If age alone—and not school experiences—accounted for

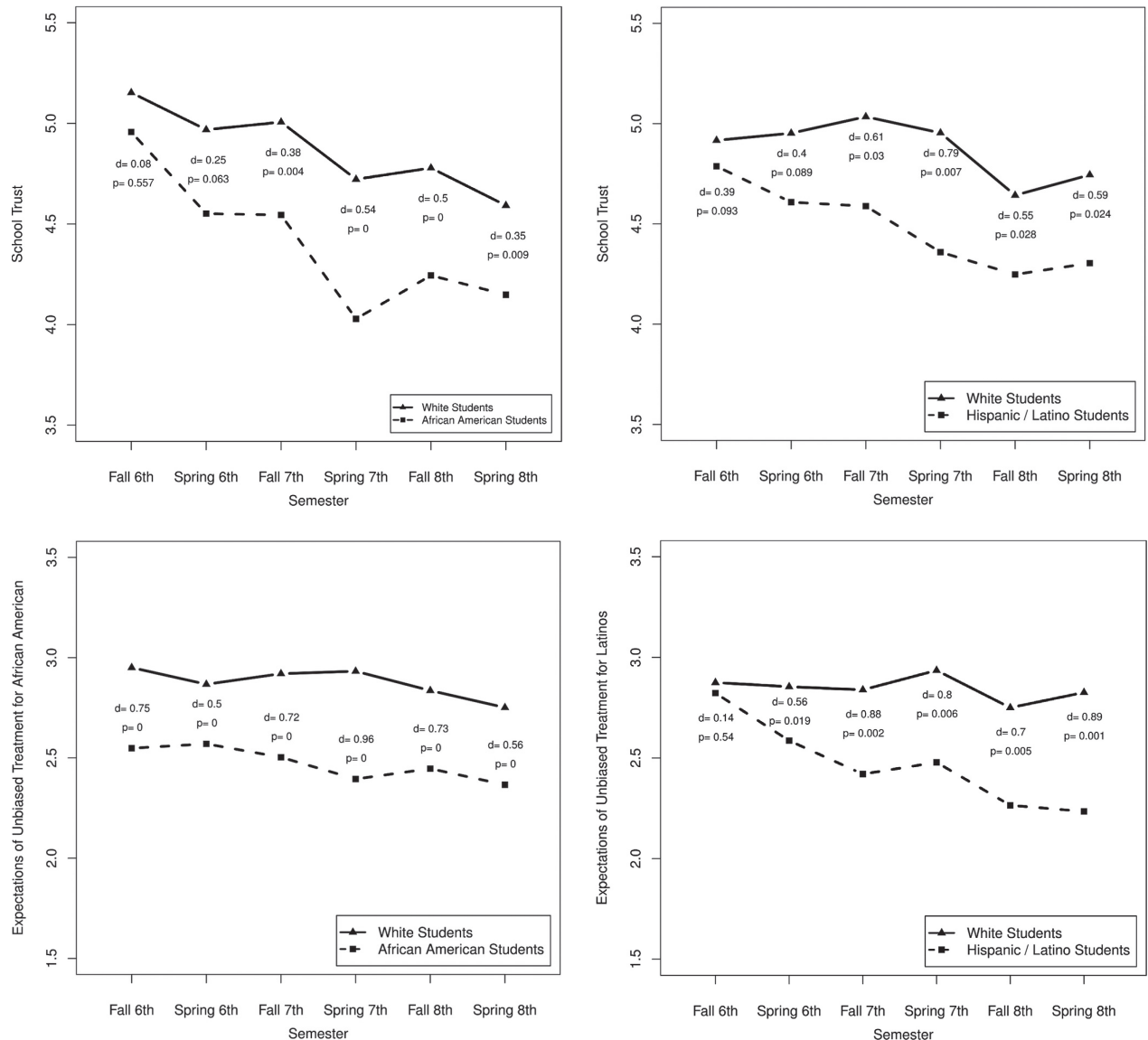


Figure 2. The racial and ethnic trust gap emerges in seventh grade and is largest in spring of seventh grade in Studies 1 and 2. Note. Values are covariate adjusted, controlling for premiddle school achievement, gender, and cohort. *p*-values are from *t*-tests; *d* = Cohen's *d*. Results are from independent regression models.

the emergence of the trust gap across the middle school years, then one might expect the decline in trust over middle school to be a monotonically decreasing linear function. However, when inspecting Figure 2, the change in trust in the summer is mostly flat or positive, whereas the change in trust during the school year mostly slopes downward. This is consistent with the explanation that loss of trust emerges from experiences in school not processes tied to chronological age or pubertal maturation alone.

To test this formally, in another growth model conducted with all students, we estimated a "school year" slope and a "summer" slope by fixing weights on two latent slope variables ("School year:" fall sixth grade = 0, spring sixth grade and fall seventh grade = 1, spring seventh grade and fall eighth grade = 2, spring eighth grade = 3, "Summer:" fall sixth grade and spring sixth grade = 0, fall seventh grade and spring seventh grade = 1, and fall eighth grade and spring eighth grade = 2). The model fit



the data acceptably:  $\chi^2(22) = 48.95$ ,  $p < .001$ ; RMSEA = .07, 90% CI [0.04, 0.09]; and CFI = .97.

For all students, "school year" trust declined significantly over time, unstandardized  $b = -0.32$ ,  $p < .001$ , 95% CI [-0.42, -0.22], and "summer" trust did not, unstandardized  $b = 0.10$ ,  $p = .092$ , 95% CI [-0.02, 0.21]. A Wald test of parameter constraints showed that the "school year" decline was steeper than the "summer" decline,  $W = 17.01$ ,  $df = 1$ ,  $p < .001$ .

Moreover, student race significantly predicted the latent "school year" slope, unstandardized  $b = -0.10$ ,  $p = .036$ , 95% CI [-0.18, -0.01]; model fit:  $\chi^2(26) = 54.39$ ,  $p < .001$ ; RMSEA = .06, 90% CI [0.04, 0.09]; CFI = .96, and not predict the summer slope ( $p = .088$ ). In sum, African American students lost more trust during the school year than their White peers. They did not lose more trust over the summer.

*Which school experiences might account for the growth of the trust gap?*

*Procedural injustice in school discipline.* Students may discern whether to trust the school based on the social reality of whether suspicion and punishments primarily target members of their racial and ethnic minority groups. Students' private grades may matter as well (which is why we control for them), but discipline-related events, such as a teacher sending a student to the principal, are more readily observable to all children.

Official records revealed that African American students received significantly more discipline throughout middle school. This was true at every grade level in the continuous data and when inspecting a dichotomous measure of whether a student had any incident at all (sixth grade: White = 24%, African American = 55%; seventh grade: White = 40%, African American = 67%; eighth grade: White = 41%; African American = 69%). Although disciplinary events rose with time for both racial groups, the largest race gap was apparent in sixth grade.

This racial group difference does not appear to be driven by differences in social class or in academic achievement. In a Poisson regression predicting number of discipline incidents in sixth to eighth grades with student race and relevant covariates (see above) a substantial student race gap in discipline remained, unstandardized  $b = 0.90$  events,  $t(269) = 13.78$ ,  $p < .001$ , 95% CI [0.77, 1.03], and this was relatively undiminished compared with the race gap with no covariates (raw gap: 5.8 incidents, covariate-adjusted gap: 4.16 incidents). Thus, on the basis of the allocation of punishment in this school,

African American students may have wondered whether the rules of the institution were being applied fairly to their group.

This is especially apparent when limiting analyses to what we call "judgment call" incidents—operationally, "defiance" and "disobedience," as coded by the school. In much prior research, there has been no group disparity in discipline for more objectively apparent infractions—such as bringing a weapon to school—but African American students were far more likely to be disciplined for subjective infractions such as disrespect, disobedience, loitering, or excessive noise (Gregory & Weinstein, 2008; Skiba et al., 2002). Students' levels of trust may be more affected by subjective than objective incidents. To be disciplined for an objectively apparent matter, like fighting, is difficult to quarrel with. But to be disciplined for talking out of turn, and to see fellow African Americans more frequently sentenced to such disciplinary measures, is perhaps more likely to raise questions about teachers' trustworthiness.

In the first cohort, the only cohort with school codes for incident type, African American students outnumbered their White peers roughly three to one for defiance or disobedience (25% vs. 8%, respectively), a significant difference  $\chi^2(1) = 6.52$ ,  $p = .012$ , and one that corresponds to the group difference in an analyses of all 17 middle schools in a large district (Skiba et al., 2002). For objective incidents such as cheating, fighting, or obscene gestures, the race gap was not significant, (7% vs. 4%),  $\chi^2(1) = 0.56$ ,  $p = .453$ .

*Awareness of bias in enforcement of school policies.* In addition to receiving more discipline, African American students also *perceived* bias in disciplinary sentencing (see Figure 2). Analyses of the awareness of bias survey items showed that African American students were more likely to expect that they, and not their White peers, would be disciplined for the same events (e.g., being in the hallway without a hall pass), at every measurement occasion,  $ps < .05$ . White students, by contrast, saw no bias in discipline. At every measurement occasion except the last one, 80% of White students expected equal treatment for White and African American students (Figure 2). Only at the end of eighth grade did White students show a significant tendency to see bias against African Americans, but they still saw less bias than African American students did at the beginning of sixth grade, as shown in Figure 2. By contrast, for African American students, fewer than 55% expected equal treatment at every wave except the first month of sixth grade.

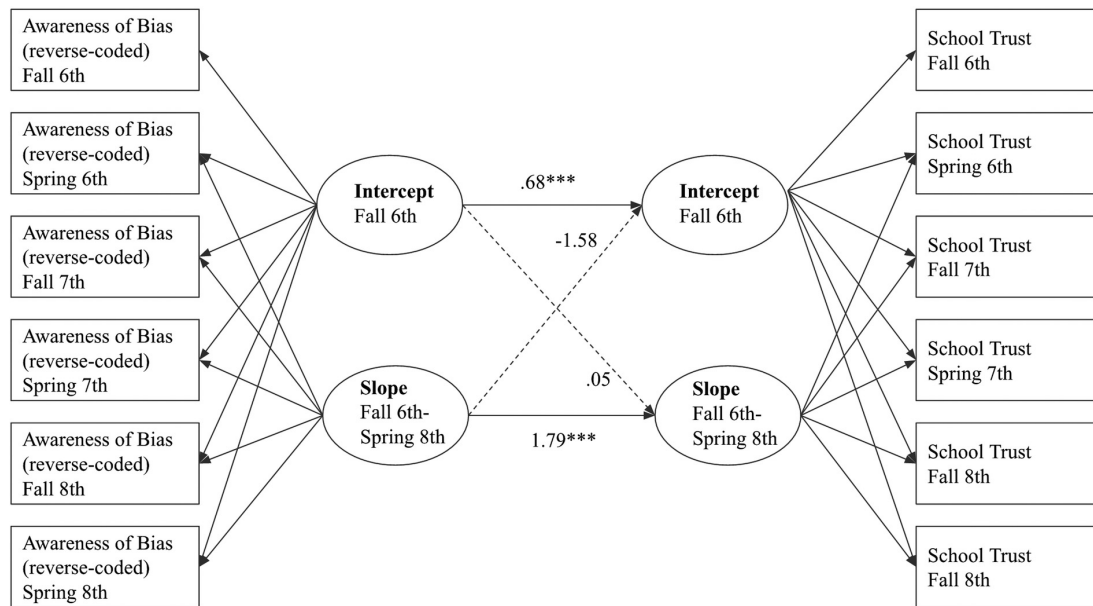


Figure 3. Recursive processes in awareness of bias and school trust among African American middle school students: a parallel growth curve analysis in Study 1.

Note. Unstandardized coefficients. Covariates were gender, prior achievement (fifth-grade grade point average and fall sixth-grade test scores), and cohort. "Awareness of bias" is coded so that higher values correspond to greater expectations of equal treatment (i.e., lower awareness of bias), matching Figure 2, so that all correlations with trust were expected to be positive. \*\*\* $p < .001$ .

*Recursive process.* A parallel growth curve model explored the relationship between African American students' awareness of bias in disciplinary decisions and their level of trust. The model estimated a latent intercept at fall of sixth grade as well as a latent slope from that point until spring of eighth grade, separately for both awareness of bias and trust (Figure 3).

We found evidence consistent with the recursive process displayed in Figure 1. African American students who reported greater awareness of disciplinary bias (reverse-coded) in the fall of sixth grade also reported lower trust at the same time point (i.e., the intercepts were correlated), unstandardized  $b = 0.68$ ,  $t(132) = 3.97$ ,  $p < .001$ , 95% CI [0.28, 1.054]. Next, a change in awareness of bias (reverse-coded) predicted a concurrent change in trust (i.e., the slopes were correlated),  $b = 1.79$ ,  $t(132) = 7.58$ ,  $p < .001$ , 95% CI [0.96, 2.61] (these results were no different when estimating the "school year" slopes only). These correlational data do not isolate which of the two constructs, mistrust or awareness of disciplinary bias, have temporal causal precedence, but rather demonstrate their pattern of reciprocal reinforcement through time.

Furthermore, if perceived bias and mistrust reinforced each other, then perceived bias should increase with time. Indeed, the growth curve model

found a significant slope for awareness of bias (not reverse-coded), unstandardized  $b = 0.06$ ,  $p = .002$ , showing that African American adolescents perceived increasingly high levels of bias through middle school. This was in spite of the fact, noted previously, that the objective size of the gap in disciplinary sentencing, if anything, decreased with time (although disciplinary events rose for all students, regardless of race, over time). In sum, students who were more aware of bias reported lower trust, and students who became more aware of bias lost more trust—and vice versa.

This process did not appear to be driven by students who were the most likely to be disciplined. Controlling for students' fifth-grade discipline incidents did not change the significance levels of either of these results. Moreover, there was no interaction involving fifth-grade discipline (in a multiple-group analysis), suggesting that the patterns held for students with a history of disciplinary problems and those without one. Furthermore, African American students showed a significant decline in trust (i.e., a significant slope) even when confined to those who *never* received a disciplinary infraction during middle school. These supplemental analyses dovetail with past research suggesting that vicarious awareness of procedural injustice can affect institutional trust, even when one has not

been the direct recipient of unfair treatment (Fagan & Tyler, 2005).

#### *Does a Loss of Trust Predict Future Discipline Incidents?*

What implications might the loss of school trust have for future behavior? A series of regression analyses found that lower trust each spring was associated with disciplinary behavior the following year. Poisson regressions predicting discipline incidents showed that lower trust in the spring of sixth and seventh grades predicted African American students' discipline incidents in the subsequent year: spring sixth-grade trust predicting seventh-grade discipline,  $b = -0.26$ ,  $t(109) = -5.53$ ,  $p < .001$ , 95% CI  $[-0.35, -0.17]$ ; spring seventh-grade trust predicting eighth grade discipline,  $b = -0.22$ ,  $t(102) = -4.172$ ,  $p < .001$ , 95% CI  $[-0.32, -0.12]$ . Hence, one year's level of institutional trust predicted the next year's level of behavioral defiance of institutional policies.

#### *Long-Term Consequences of Loss of Trust*

Did a loss in trust predict long-term developmental outcomes? We estimated residual scores from regressions of spring trust on fall trust, for each of the three school years: sixth, seventh, and eighth. Lower values correspond to a greater loss of trust. The advantage of the residual score is that it is independent of either of the two scores used to compute it (supplementary analyses using raw change scores yielded the same substantive results).

Seventh-grade residual trust scores predicted on time enrollment at a 4-year college for African American students, standardized  $\beta = .19$ ,  $Z = 2.26$ ,  $p = .024$ , 95% CI  $[0.025, 0.36]$ , meaning that students who lost more trust than would be expected in seventh grade were less likely to end up at a 4-year college. (Similar results emerged in a supplementary ordinal regression analysis that examined college enrollment at three levels, 0 = no college, 1 = 2-year college, 2 = 4-year college,  $Z = 2.38$ ,  $p = .018$ .) Fall eighth-grade absolute levels of trust predicted college enrollment as well ( $p < .01$ ), and no other absolute or residual measure of trust did (see Supporting Information).

Thus, it appears that the damage to African American students' trust by the end of seventh grade or the beginning of eighth grade significantly predicted whether they made it to a 4-year college the year after high school. Crucially, this was true even after controlling for relevant predictors of

college enrollment such as premiddle school academic performance. In contrast, trust spring residual scores or within-year change scores in any year were not associated with college enrollment for White students (see Figure S5).

#### *Effect of the "Wise Feedback" Intervention*

One limitation with the analyses reported so far is endogeneity. Students' social realities in their institutions, their mental representations of them, and their behaviors over time occur in a mutually causal system, creating a confound for causal inference. We are fortunate, however, that it was possible to supplement a portion of the correlational analysis with an experimental one. As noted, students were randomly assigned at the individual level to receive trust-sustaining "wise feedback" from their seventh-grade teacher in the context of a potentially unfair interpersonal event: being criticized on a writing assignment (recall that, in a previous article reporting on these data, the note sustained African American students' trust at the end of seventh grade; Yeager et al., 2014, Studies 1–2).

*Eighth-grade discipline.* Receiving a trust-sustaining "wise feedback" intervention in the spring of seventh grade significantly reduced African American students' eighth-grade discipline incidents,  $t(43) = 2.96$ ,  $p = .005$ ,  $d = .67$ . There was no effect for White students ( $p = .75$ ), and the Race  $\times$  Condition interaction was significant,  $t(81) = 2.04$ ,  $p = .045$ . Figure 4 reports raw and covariate-adjusted means. In the control condition, African American students received four times more disciplinary citations as compared to White students. In the treatment group that gap was halved. These results represent a conceptual replication of research that has changed teachers' chronic behaviors, improving students' feelings that teachers respected them, thus reducing race-related discipline disparities (Gregory et al., 2016; Okonofua, Paunesku, et al., 2016).

These results support the theory presented in Figure 1, which contends that adolescents engage in inductive reasoning on the basis of interpersonal interactions but, after forming a mental model, are also partly top-down in their assessments of future attributionally ambiguous interactions. Here, a trust-restoring experience with a single teacher in seventh grade affected discipline in eighth grade, even as students entered into new interactions with different teachers and authority figures in the school (for analogous year-over-year

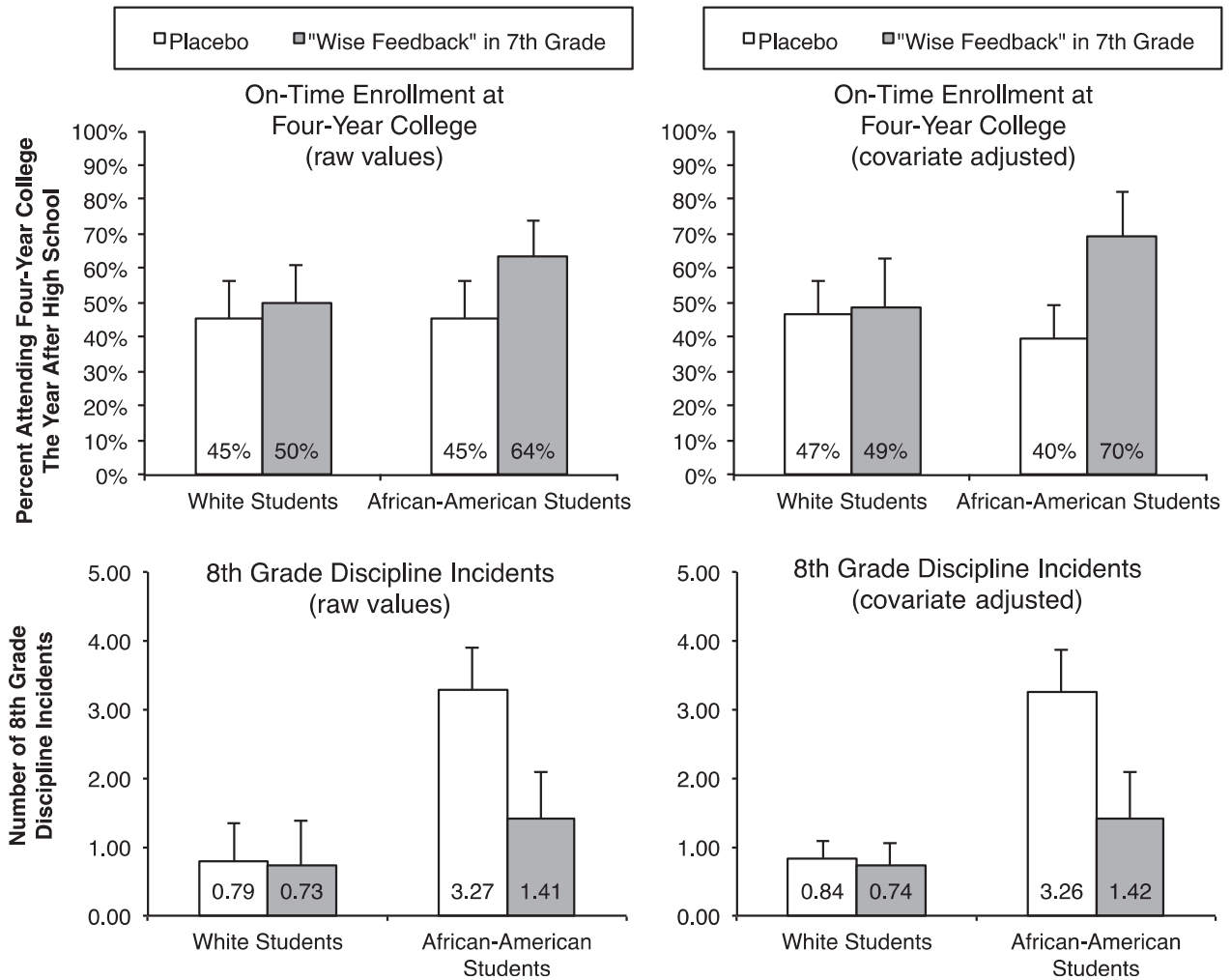


Figure 4. Longitudinal effects of trust-restoring "wise feedback" intervention (administered spring of seventh grade) on eighth-grade discipline incidents and on time enrollment in a 4-year college 5½ years postintervention, in Study 1.

Note. Covariates include prior achievement (grade point average and test scores), preintervention trust, gender, and cohort. Covariate-adjusted values for each racial group estimated at the mean for each racial group.

effects, see Gregory et al., 2016). This suggests that it is possible to interrupt or at least slow the top-down appraisals that contribute to worsening mistrust.

**College enrollment.** Next, the wise feedback intervention in spring of seventh grade—the period when we found that changes in trust were most predictive of college outcomes—significantly increased African American students' likelihood of attending a 4-year college the year immediately following high school, 5.5 years posttreatment,  $Z = 2.58$ ,  $p = .010$  (Covariate-adjusted values: Control = 40%, Wise feedback = 70%, see Figure 4 for raw values). Gender, prior achievement, and prior trust all significantly predicted college enrollment, but cohort did not; removing the latter did not

change the significance of the treatment effect ( $p = .01$ ). Treatment effects for White students were nonsignificant ( $p < .5$ ).

There were no treatment effects on whether adolescents enrolled at any college at all (2 or 4 years),  $Z = -0.16$ ,  $p = .880$ , only on 4-year college enrollment. In a supplemental ordinal regression analysis when 0 = no college enrollment, 1 = 2-year college, and 2 = 4-year college, the treatment effect was marginally significant,  $Z = 1.69$ ,  $p = .092$ .

Recall, however, that this experiment was conducted among the subset of "B" and "C" average students who were already more likely to attend college than students with lower grades, and so there was range restriction. Within the subset of those who attended any college at all ( $N = 37$  of



44), we analyzed treatment effects on 4-year college enrollment (1 = 4-year college, 0 = 2-year college, NA = not enrolled in college). In that analysis, effects of the wise feedback treatment were stronger,  $Z = 2.95$ ,  $p = .003$ , with covariates. Thus, at least in our small sample of middle-class “B” and “C” African American students who eventually attended college, the trust-sustaining “wise-feedback” intervention resulted in greater likelihood of attending a 4-year college instead of a 2-year college.

## Study 2

### Method

#### Participants and Procedures

Study 2 was a 1-year cohort sequential study that followed the entire middle school (sixth, seventh, and eighth grade) at a rural school in Colorado for 1 year ( $N = 206$ ). Students were either Hispanic or Latino (44%) or White (56%). Unlike Study 1, the negatively-stereotyped ethnic minority students were largely working class. These differences between the studies permitted us to assess the generality of our findings to a new negatively stereotyped group (Latino Americans) that was relatively economically disadvantaged and living in an altogether different geographic location from the students in Study 1.

Survey procedures were analogous to Study 1. Adolescents completed fall and spring survey measures assessing school trust and awareness of procedural injustice, identical to Study 1, except the procedural injustice in school policies survey measures (after piloting) referred to “Hispanic” students, not “African American” students. Students were surveyed only for 1 year of middle school, and because this study was conducted years after Study 1, students have not yet been tracked through college. Furthermore, discipline incidents in the year after the study were not available.

### Results

#### *Emergence of the Latino–White Trust Gap in Middle School*

As in Study 1, a significant ( $p < .05$ ) ethnic group gap in school trust emerged during fall of 7th grade and was sustained through the end of middle school (see Figure 2); once again, the trust gap was not significant in 6th grade. As in Study 1, the

largest loss in trust appeared in the 7th grade school year, was strongest for minority rather than majority students, with the largest gap in trust between the two ethnic groups occurring in the spring of 7th grade. Thus, analyses of school trust in a cohort-sequential design with Latinos replicated Study 1’s results with African Americans.

#### *What Caused the Trust Gap?*

Unlike Study 1, in the fall of sixth grade there was no initial significant Latino–white difference in expectations of bias in application of school policies (Figure 2). However, by the spring of sixth grade a significant difference emerged, and from then on the results of Study 2 paralleled Study 1.

Next, we examined the relationship between awareness of bias and changes in trust. Because Study 2 is a cohort sequential design, not a true longitudinal design like Study 1, the latent change score model is not identified, as it requires at least three observations. Therefore, we estimated cross-lagged autoregressive models with Latino participants from all three grades (sixth, seventh, and eighth) and tested the effects of fall awareness of bias on spring trust.

The cross-lagged model showed that the extent to which Latino students believed there was injustice in school discipline in the *fall* was a significant predictor of their subsequent trust of school in the *spring*, in cross-lagged models ( $b = 0.53$ ,  $p < .001$ ; see Figure 5). The reverse was not true, suggesting a path from awareness of bias to trust, and not vice versa. This path was not significantly different for any of the three grade levels, when constraining paths across grade levels in a multigroup analysis (see Supporting Information). (We have also reconducted Study 1’s analyses using cross-lagged models, and they are reported in the online supplement, Figure S6. These yield the same conclusions, with the exception that both paths, from trust to bias

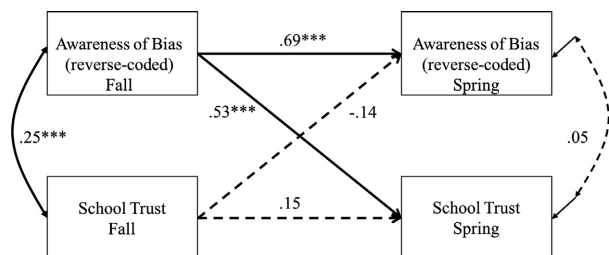


Figure 5. Cross-lagged models show that awareness of bias in discipline predicts subsequent lower levels of trust among Latinos in Study 2. Unstandardized regression coefficients. \*\*\* $p < .001$ .

and bias to trust, are significant at different waves.) These findings extend the parallel growth curve analysis in Study 1 by showing that awareness of bias in the fall predicts lower trust of school in the spring.

### Discussion

In a Pew Center survey of adults in the United States, 61% of African Americans and 53% of Latinos reported low levels of trust in the fairness of American society, as compared to only 32% of White Americans (Taylor, Funk, & Clark, 2007). Analogous gaps exist with regard to specific U.S. institutions. For instance, 70% of African Americans recently reported low levels of trust in the police compared to 43% of White Americans (Jones, 2015). Such racial and ethnic gaps in adults' levels of trust are large. They may appear unavoidable, normative, and unsurprising. A key contribution of this article is to offer a testable social and developmental psychological process model for how racial and ethnic gaps in institutional trust emerge during adolescence and affect long-term developmental outcomes (Figure 1). A second contribution is to show that such gaps and their consequences need not be inevitable.

Specifically, this research used a theory-informed developmental frame to understand the formation of institutional trust among racial and ethnic minority adolescents and to explore its behavioral implications over an 8-year period. Trust declined every semester of middle school, and this decline happened more quickly and strongly for African Americans (Study 1) and Latino (Study 2) students. Racial and ethnic minority students' decline in trust was predicted by their awareness of bias—that is, their awareness of the possibility of procedural injustice at their school. This finding was confirmed using a parallel growth curve model (Study 1) and a cross-lagged model (Study 2). These results are consistent with much prior research showing that trust is tied to perceptions of procedural justice and predicts institutional compliance (Levi, 1997; Tyler & Blader, 2003).

One intriguing aspect of the process shown in Figure 1 is its recursive nature. Once students' sense of trust or distrust was formed, it seemed to feed off its consequences, producing perceptions of procedural injustice that caused trust to decline further (also see Fagan & Tyler, 2005; Gregory & Weinstein, 2008). Moreover, that decline in trust seemed to increase the likelihood of discipline

infractions, creating the very social reality that precipitated it. These feedback loops proceed often hidden from the view of teachers and administrators because they unfold slowly and are partly psychological in nature. But their cumulative effect is a large trust gap by seventh grade that disfavored racial and ethnic minority students. Years later, the drop in trust in the transition to seventh grade and then eighth grade seemed to have lingering consequences, in the form of lower 4-year college enrollment for African Americans.

Trust, it seems, sat “in the middle” between social reality and later behavioral outcomes such as disciplinary infractions and college enrollment (cf. Olson & Dweck, 2008). We know this from an intervention in Study 1 that experimentally bolstered African Americans' sense of trust in the face of sharp criticism of their work in the seventh grade (previously shown in Yeager et al., 2014, Study 2). Because the link between trust and later outcomes depends on a continual feedback loop, an early experience that refuted the plausibility of procedural injustice had long-term effects, presumably through a kind of developmental cascade from trust to engagement and into educational pathways (cf. Masten et al., 2005).

African American seventh graders who received wise feedback on an essay—conveying that the teacher believed in their potential to reach a higher standard, thus reassuring students that they would be seen based on their merits rather than through the lens of a negative stereotype about the intellectual ability of their racial group—benefited. Our intervention was based on a body of previous theoretical and empirical research (Cohen et al., 1999; Gregory & Weinstein, 2008; Shouse, 1996) and was timed to the emergence of the trust gap (see Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009). Although the objective experience of receiving “wise feedback” was short, the psychological and developmental consequences seemed long-lasting. Adolescents receiving the note were assigned fewer disciplinary infractions later according to official records and, nearly 6 years later, were more likely to attend a 4-year college according to the NSC.

Trust, and the trust-restoring wise feedback intervention, did not predict subsequent college persistence for White students (Figure S5). Why not? For students with group-based advantages, such as majority-group students, there may simply be greater redundancy in the system. When one's group is positively stereotyped and overrepresented in a domain, a loss of trust or a poor relationship with a teacher might be a temporary setback but

not fatal. For a stigmatized student, however, outcomes might depend more on the accident of having a teacher who believes in them (for an analogous example from low-income children in early childhood, see Tucker-Drob, 2012). Thus, although the present results offer an exciting possibility for what might be done to prevent race-based achievement gaps, they simultaneously act as an indictment of the educational system, by demonstrating that minority students' outcomes are more contingent on everyday experiences of respect—experiences that may occur too rarely and by happenstance.

### *Limitations and Future Directions*

The wise feedback experiment has limitations. First, the evidence is based on only one school. Second, the sample size was small, and small sample sizes can lead to false-positive findings or to overstated effect sizes when researchers have flexibility in data analysis. Here, sample size was constrained by the number of B and C students in the two cohorts, and analysis was constrained to be parallel across types of analysis—correlational and experimental—which yielded similar results. In addition, the results are robust across many (but not all) alternative model specifications (see both significant and nonsignificant model specifications at [osf.io/3hpu8](https://osf.io/3hpu8)).

At the same time, the overall sample size was approximately 70% of the longitudinal sample for the Perry Preschool project, which has had a tremendous influence on early childhood policy (e.g., Heckman, 2006). Furthermore, unlike the Perry Preschool project, in this study there were no problems with randomization or missing data (because random assignment was effective, Yeager et al., 2014; and outcome data were collected from the NSC). We also note that prior psychological interventions that initially involved a few dozen treated students (e.g., Aronson, Fried, & Good, 2002) later showed reproducible findings when conducted with many thousands of students (Paunesku et al., 2015) and when analyses were preregistered and data were collected by third-party research firms (Yeager, Romero, et al., 2016).

Nevertheless, it will be important for future experiments to reproduce our method with a larger sample and in more heterogeneous contexts. Doing this will not be without challenges. The present experiments require working closely with schools over several years to guide teachers in designing assignments and hand-writing personal notes to students, and tracking these students over the

ensuing 6 years. Moreover, we know very little about what contextual conditions might moderate the effects of wise feedback, but we imagine they are plentiful.

Although we have focused on the middle school transition, the processes documented here did not begin in sixth grade and do not stop in eighth grade. It will be critical in future research to understand how to prevent experiences of racial and ethnic discrimination that begin early in childhood and also to develop methods to regain racial and ethnic minority students' trust after they transition to high school. For instance, one possibility is that students reengage their potential to trust when they transition to a novel institution—asking themselves whether the “old codes” apply or not (cf. Coates, 2015). Perhaps timely intervention at later life transitions might approximate the effects of the wise feedback intervention documented here (for an example among college students, see Yeager, Walton, et al., 2016).

We emphasize that our theoretical model does not “psychologize” the social problem of racial disparities in discipline and college enrollment. Nor does it “sociologize” it and ascribe outcomes solely to unalterable structural causes. It is, as we show in Figure 1, the interaction of structure with psyche that drives outcomes over time (Cohen, Garcia, & Goyer, 2017). School presents a social reality to which all students, especially negatively stereotyped minority groups, must psychologically adapt. That psychological adaptation, in turn, can reinforce the social reality, as when African American middle school students act out against a system they perceive as unjust. It is neither the attributes of the child, such as a troublesome nature or a behavioral disorder, nor the social environment alone that is the driver of inequality. Rather it is the unfolding transaction between the child and the environment (Cohen et al., 2017; Gregory & Weinstein, 2008; Okonofua, Walton, et al., 2016).

Our approach echoes Bronfenbrenner's (1977) early formulation of ecological systems theory and aligns with contemporary social-cognitive theories of development (Olson & Dweck, 2008). Mental representations, the product of observing and interacting with a social context, continue to exert an influence on development for years to come. Yet by the logic of causal mediation, intervening in a timely way slowed the psychological accretions of injustice and lessened its lasting harm.

Our theoretical perspective is consistent with phenomenological variants on ecological systems theories, which address the social, structural,

cultural, and historical contexts in which youth develop (Spencer, 2006). A mental model of trust can act as a protective factor or a risk factor. Trust allows a person to reap the rewards of engagement with the system, but it also puts them at risk of being taken advantage of. Trust, absent justice, might be misplaced, but well-placed trust can help people thrive. Hence, institutions would do well to simultaneously improve fairness and reduce bias while also addressing the psychological legacy of previous injustices. Said differently, wise feedback is not, on its own, the solution to race gaps in discipline. Schools and teachers need to be trustworthy. When they are, then wise feedback can help clear the attributional air and redirect students' working mental models of the institution, so that students can profit from the relationships and instruction in their schools.

These data inform education policy and practice. By experimentally testing the wise feedback intervention, it demonstrates a method for helping teachers create a classroom climate that is more likely to maintain the trust of students who may contend with discrimination (also see Gregory et al., 2016; Okonofua, Paunesku, et al., 2016). Based on prior correlational (Gregory & Weinstein, 2008; Shouse, 1996) and experimental research (Cohen & Steele, 2002; Cohen et al., 2009; Yeager et al., 2014), the intervention was not only precisely targeted to address the key psychological process but was timed to intervene on that process at a crucial stage: when mistrust was likely to emerge and to exert a growing impact. This is because a process can be easier to affect at its beginning. Once it has accumulated consequences—the child is labeled as a “troublemaker,” sees his permanent record tainted by a suspension—the process will have a momentum that is much harder to halt (Okonofua, Paunesku, et al., 2016; Okonofua, Walton, et al., 2016).

Of course, truly “wise” educators do not simply append notes to essays and end their interventions there. Instead, they continually send the message that their students are capable, valued, and respected, weaving it into the culture of the classroom. Our studies demonstrate the way that larger cultural forces infuse the interactions between teachers and students. How to break free of their influence is a craft that requires both wisdom and tact.

## References

- Aronson, J. M., Fried, C. B., & Good, C. (2002). Reducing the effects of stereotype threat on African American college students by shaping theories of intelligence. *Journal of Experimental Social Psychology, 38*, 113–125. doi:10.1006/jesp.2001.1491
- Baumrind, D. (1997). The discipline encounter: Contemporary issues. *Aggression and Violent Behavior, 2*, 321–335. doi:10.1016/S1359-1789(97)00018-9
- Bigler, R. S., Averhart, C. J., & Liben, L. S. (2003). Race and the workforce: Occupational status, aspirations, and stereotyping among African American children. *Developmental Psychology, 39*, 572–580. doi:10.1037/0012-1649.39.3.572
- Bigler, R. S., & Liben, L. S. (2006). A developmental intergroup theory of social stereotypes and prejudice. *Advances in Child Development and Behavior, 34*, 39–89. doi:10.1016/S0065-2407(06)80004-2
- Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist, 32*, 513–531. doi:10.1037/0003-066X.32.7.513
- Brown, C. S., & Bigler, R. S. (2005). Children's perceptions of discrimination: A developmental model. *Child Development, 76*, 533–553. doi:10.1111/j.1467-8624.2005.00862.x
- Bryk, A. S., & Schneider, B. (2002). *Trust in schools: A core resource for improvement*. New York, NY: Russell Sage Foundation.
- Coates, T.-N. (2015). *Between the world and me*. New York, NY: Random House Publishing Group.
- Cohen, G. L., & Garcia, J. (2005). “I am us”: Negative stereotypes as collective threats. *Journal of Personality and Social Psychology, 89*, 566–582. doi:10.1037/0022-3514.89.4.566
- Cohen, G. L., Garcia, J., & Goyer, J. P. (in press). Turning point: Targeted, tailored, and timely psychological intervention. In A. Elliot, C. S. Dweck & D. S. Yeager (Eds.) *Handbook of competence and motivation (2nd ed.)*: Theory and application. New York, NY: Guilford.
- Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., & Brzustoski, P. (2009). Recursive processes in self-affirmation: Intervening to close the minority achievement gap. *Science, 324*, 400–403. doi:10.1126/science.1170769
- Cohen, G. L., & Steele, C. M. (2002). A barrier of mistrust: How negative stereotypes affect cross-race mentoring. In J. M. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (pp. 303–327). Bingley, UK: Emerald Group.
- Cohen, G. L., Steele, C. M., & Ross, L. D. (1999). The mentor's dilemma: Providing critical feedback across the racial divide. *Personality and Social Psychology Bulletin, 25*, 1302–1318. doi:10.1177/0146167299258011
- Coopersmith, J. (2009). *Characteristics of public, private, and bureau of indian education elementary and secondary school teachers in the United States: Results from the 2007–08 schools and staffing survey*. First Look. NCES 2009-324. Washington D.C.: National Center for Education Statistics.
- Dynarski, S. M., Hemelt, S. W., & Hyman, J. M. (2015). The missing manual: Using National Student Clearinghouse data to track postsecondary outcomes. *Educational Evaluation and Policy Analysis, 37*, 53s–79s. doi:10.3102/0162373715576078



- Fagan, J., & Tyler, T. R. (2005). Legal socialization of children and adolescents. *Social Justice Research, 18*, 217–241. doi:10.1007/s11211-005-6823-3
- Gambetta, D. (1988). Can we trust trust? In D. Gambetta (Ed.), *Trust: Making and breaking cooperative relations* (pp. 213–235). New York, NY: Basil Blackwell.
- Goffman, E. (1963). *Stigma: Notes on the management of spoiled identity*. New York, NY: Simon & Schuster.
- Gregory, A., Hafen, C. A., Ruzek, E., Mikami, A. Y., Allen, J. P., & Pianta, R. C. (2016). Closing the racial discipline gap in classrooms by changing teacher practice. *School Psychology Review, 45*, 171–191.
- Gregory, A., & Weinstein, R. S. (2004). Connection and regulation at home and in school predicting growth in achievement for adolescents. *Journal of Adolescent Research, 19*, 405–427. doi:10.1177/0743558403258859
- Gregory, A., & Weinstein, R. S. (2008). The discipline gap and African Americans: Defiance or cooperation in the high school classroom. *Journal of School Psychology, 46*, 455–475. doi:10.1016/j.jsp.2007.09.001
- Harber, K. D. (1998). Feedback to minorities: Evidence of a positive bias. *Journal of Personality and Social Psychology, 74*, 622–628. doi:10.1037/0022-3514.74.3.622
- Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science, 312*, 1900–1902. doi:10.1126/science.1128898
- Irvine, J. J. (2003). *Educating teachers for diversity: Seeing with a cultural eye*. New York, NY: Teachers College Press.
- Jones, J. M. (2015, June 19). In U.S., confidence in police lowest in 22 years. Omaha, NE: Gallup. Retrieved from <http://www.gallup.com/poll/183704/confidence-police-lowest-years.aspx>
- Killen, M., Mulvey, K. L., & Hitti, A. (2013). Social exclusion in childhood: A developmental intergroup perspective. *Child Development, 84*, 772–790. doi:10.1111/cdev.12012
- Laupa, M., Turiel, E., & Cowan, P. A. (1995). Obedience to authority in children and adults. In M. Killen & D. Hart (Eds.), *Morality in everyday life: Developmental perspectives* (pp. 131–165). Cambridge, UK: Cambridge University Press.
- Lee, K. H., Siegle, G. J., Dahl, R. E., Hooley, J. M., & Silk, J. S. (2014). Neural responses to maternal criticism in healthy youth. *Social Cognitive and Affective Neuroscience, 10*, 902–912. doi:10.1093/scan/nsu133
- Lepper, M. R., & Woolverton, M. (2002). The wisdom of practice: Lessons learned from the study of highly effective tutors. In J. M. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (pp. 135–158). San Diego, CA: Academic Press.
- Levi, M. (1997). *Consent, dissent, and patriotism*. New York, NY: Cambridge University Press.
- Levi, M., Sacks, A., & Tyler, T. (2009). Conceptualizing legitimacy, measuring legitimating beliefs. *American Behavioral Scientist, 53*, 354–375. doi:10.1177/0002764209338797
- Lewin, K., Lippitt, R., & White, R. K. (1939). Patterns of aggressive behavior in experimentally created social climates. *The Journal of Social Psychology, 10*, 269–299. doi:10.1080/00224545.1939.9713366
- Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology, 37*, 2098–2109. doi:10.1037/0022-3514.37.11.2098
- Masten, A. S., Roisman, G. I., Long, J. D., Burt, K. B., Obradović, J., Riley, J. R., . . . Tellegen, A. (2005). Developmental cascades: Linking academic achievement and externalizing and internalizing symptoms over 20 years. *Developmental Psychology, 41*, 733–746. doi:10.1037/0012-1649.41.5.733
- McKown, C. (2013). Social equity theory and racial-ethnic achievement gaps. *Child Development, 84*, 1120–1136. doi:10.1111/cdev.12033
- McKown, C., & Weinstein, R. S. (2003). The development and consequences of stereotype consciousness in middle childhood. *Child Development, 74*, 498–515.
- Muthén, L. K., & Muthén, B. O. (1998–2011). *Mplus user's guide*. Sixth Edition. Los Angeles, CA: Author.
- Okonofua, J. A., & Eberhardt, J. L. (2015). Two strikes: Race and the disciplining of young students. *Psychological Science, 26*, 617–624. doi:10.1177/0956797615570365
- Okonofua, J. A., Paunesku, D., & Walton, G. M. (2016). A brief intervention to encourage empathic discipline cuts suspension rates in half among adolescents. *Proceedings of the National Academy of Sciences of the United States of America, 113*, 5221–5226.
- Okonofua, J. A., Walton, G. M., & Eberhardt, J. L. (2016). A vicious cycle: A social-psychological account of extreme racial disparities in school discipline. *Perspectives on Psychological Science, 11*, 381–398.
- Olson, K. R., & Dweck, C. S. (2008). A blueprint for social cognitive development. *Perspectives on Psychological Science, 3*, 193–202. doi:10.1111/j.1745-6916.2006.00007.x
- Paunesku, D., Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S., & Dweck, C. S. (2015). Mindset interventions are a scalable treatment for academic underachievement. *Psychological Science, 26*, 284–293. doi:10.1177/0956797615571017
- Peper, J. S., & Dahl, R. E. (2013). The teenage brain: Surging hormones—brain-behavior interactions during puberty. *Current Directions in Psychological Science, 22*, 134–139.
- President's Task Force on 21st Century Policing. (2015). *Final report of The President's Task Force on 21st Century Policing*. Washington, DC: Office of Community Oriented Policing Services.
- Purdie-Vaughns, V., Steele, C. M., Davies, P. G., Dittmann, R., & Crosby, J. R. (2008). Social identity contingencies: How diversity cues signal threat or safety for African Americans in mainstream institutions. *Journal of Personality and Social Psychology, 94*, 615–630. doi:10.1037/0022-3514.94.4.615

- Roberts, B. W., & Caspi, A. (2003). The cumulative continuity model of personality development: Striking a balance between continuity and change in personality traits across the life course. In U. M. Staudinger & U. E. R. Lindenberger (Eds.), *Understanding human development: Lifespan psychology in exchange with other disciplines* (pp. 183–214). Dordrecht, NL: Kluwer Academic.
- Ruble, D. N., Alvarez, J., Bachman, M., Cameron, J., Fuligni, A., Garcia Coll, C., & Rhee, E. (2004). The development of a sense of “we”: The emergence and implications of children’s collective identity. In M. Bennett & F. Sani (Eds.), *The development of the social self* (pp. 29–76). New York, NY: Psychology Press.
- Sagar, H. A., & Schofield, J. W. (1980). Racial and behavioral cues in black and white children’s perceptions of ambiguously aggressive acts. *Journal of Personality and Social Psychology*, 39, 590–598.
- Sellers, R. M., Copeland-Linder, N., Martin, P. P., & Lewis, R. L. (2006). Racial identity matters: The relationship between racial discrimination and psychological functioning in African American adolescents. *Journal of Research on Adolescence*, 16, 187–216.
- Shouse, R. C. (1996). Academic press and sense of community: Conflict, congruence, and implications for student achievement. *Social Psychology of Education*, 1(1), 47–68.
- Skiba, R. J., Michael, R. S., Nardo, A. C., & Peterson, R. L. (2002). The color of discipline: Sources of racial and gender disproportionality in school punishment. *The Urban Review*, 34, 317–342. doi:10.1023/A:1021320817372
- Spencer, M. B. (2006). Phenomenology and ecological systems theory: Development of diverse groups. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology, Vol. 1: Theoretical models of human development* (6th ed., pp. 829–893). New York, NY: Wiley.
- Swanson, D. P., Cunningham, M., Youngblood, J., & Spencer, M. B. (2009). Racial identity development during childhood. In H. A. Neville, B. M. Tynes, & S. O. Utsey (Eds.), *Handbook of African American psychology* (pp. 269–281). Thousand Oaks, CA: Sage.
- Taylor, P., Funk, C., & Clark, A. (2007). *Americans and social trust: Who, where and why*. Washington, DC: Pew Research Center.
- Terburg, D., & van Honk, J. (2013). Approach–avoidance versus dominance–submissiveness: A multilevel neural framework on how testosterone promotes social status. *Emotion Review*, 5, 296–302.
- Treisman, U. (1992). Studying students studying calculus: A look at the lives of minority mathematics students in college. *The College Mathematics Journal*, 23, 362–372. doi:10.2307/2686410
- Tucker-Drob, E. M. (2012). Preschools reduce early academic-achievement gaps: A longitudinal twin approach. *Psychological Science*, 23, 310–319. doi:10.1177/0956797611426728
- Tyler, T. R. (2006). Restorative justice and procedural justice: Dealing with rule breaking. *Journal of Social Issues*, 62, 307–326. doi:10.1111/j.1540-4560.2006.00452.x
- Tyler, T. R., & Blader, S. L. (2003). The group engagement model: Procedural justice, social identity, and cooperative behavior. *Personality and Social Psychology Review*, 7, 349–361. doi:10.1207/S15327957PSPR0704\_07
- Tyler, T. R., Fagan, J., & Geller, A. (2014). Street stops and police legitimacy: Teachable moments in young urban men’s legal socialization. *Journal of Empirical Legal Studies*, 11, 751–785. doi:10.1111/jels.12055
- Tyler, T. R., Goff, P. A., & MacCoun, R. J. (2015). The impact of psychological science on policing in the united states procedural justice, legitimacy, and effective law enforcement. *Psychological Science in the Public Interest*, 16, 75–109. doi:10.1177/1529100615617791
- Vanderbilt, K. E., Liu, D., & Heyman, G. D. (2011). The development of distrust. *Child Development*, 82, 1372–1380. doi:10.1111/j.1467-8624.2011.01629.x
- Vasquez, J. A. (1988). Contexts of learning for minority students. *The Educational Forum*, 52, 243–253. doi:10.1080/00131728809335490
- Yeager, D. S., Purdie-Vaughns, V., Garcia, J., Apfel, N., Brzustoski, P., Master, A., . . . Cohen, G. L. (2014). Breaking the cycle of mistrust: Wise interventions to provide critical feedback across the racial divide. *Journal of Experimental Psychology: General*, 143, 804–824. doi:10.1037/a0033906
- Yeager, D. S., Romero, C., Paunesku, D., Hulleman, C. S., Schneider, B., Hinojosa, C., . . . Dweck, C. S. (2016). Using design thinking to improve psychological interventions: The case of the growth mindset during the transition to high school. *Journal of Educational Psychology*, 108, 374–391. doi:10.1037/edu0000098
- Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They’re not magic. *Review of Educational Research*, 81, 267–301. doi:10.3102/0034654311405999
- Yeager, D. S., Walton, G. M., Brady, S. T., Akcinar, E. N., Paunesku, D., Keane, L., . . . Dweck, C. S. (2016). Teaching a lay theory before college narrows achievement gaps at scale. *Proceedings of the National Academy of Sciences of the United States of America*, 113, E3341–E3348. doi:10.1073/pnas.1524360113

### Supporting Information

Additional supporting information may be found in the online version of this article at the publisher’s website:

**Figure S1.** Exploratory Mediation Analysis of Effects of Seventh-Grade Trust (Measured, Top Panel; or Manipulated, Bottom Panel) on on time 4-Year College Enrollment, Via Eighth-Grade Behaviors, for African American Participants Only

**Figure S2.** Cross-Lagged Models Show That Awareness of Bias Predicts Subsequent Trust for

Seventh- and Eighth-Grade Hispanic/Latino Students, Controlling for Gender

**Figure S3.** Individual Differences in Awareness of Anti-African American Procedural Injustice in School Policies Do Not Predict the Onset of Overall Mental Models of Trust of School Among White Middle School Students in Study 1

**Figure S4.** The “Trust Residual Predicts College Enrollment” Model, Excluding Treated Students

**Figure S5.** The “Trust Residual Predicts College Enrollment” Model, White Students Only

**Figure S6.** Cross-Lagged Analyses Among

African American Participants in Study 1, Showing the Reciprocal Relations Between Perceived Bias and Loss of Trust Across the Middle School Years

**Table S1.** Study 1 Racial Group Differences in School Trust

**Table S2.** Study 2 Ethnic Group Differences in School Trust

**Table S3.** Descriptive Statistics for Study Variables (Study 1)

**Table S4.** Descriptive Statistics for Study Variables (Study 2)