



Experimental Effects of “Achievement Gap” News Reporting on Viewers’ Racial Stereotypes, Inequality Explanations, and Inequality Prioritization

David M. Quinn

University of Southern California

The “achievement gap” has long dominated mainstream conversations about race and education. Some scholars warn that the discourse around racial gaps perpetuates stereotypes and promotes the adoption of deficit-based explanations that fail to appreciate the role of structural inequities. I investigate through three randomized experiments. Results indicate that a TV news story about racial achievement gaps (versus a control or counter-stereotypical video) led viewers to express more exaggerated stereotypes of Black Americans as lacking education (study 1: $ES=.30$ SD; study 2: $ES=.38$ SD) and may have increased viewers’ implicit stereotyping of Black students as less competent than White students (study 1: $ES=.22$ SD; study 2: $ES=.12$ SD, n.s.). The video did not affect viewers’ explicit competence-related racial stereotyping, the explanations they gave for achievement inequalities, or their prioritization of ending achievement inequalities. After two weeks, the effect on stereotype exaggeration faded. Future research should probe how we can most productively frame educational inequality by race.

VERSION: June 2020

Suggested citation: Quinn, David M.. (2020). Experimental Effects of “Achievement Gap” News Reporting on Viewers’ Racial Stereotypes, Inequality Explanations, and Inequality Prioritization. (EdWorkingPaper: 20-237). Retrieved from Annenberg Institute at Brown University: <https://doi.org/10.26300/075t-ed79>

**Experimental Effects of “Achievement Gap” News Reporting on Viewers’ Racial
Stereotypes, Inequality Explanations, and Inequality Prioritization**

Forthcoming in *Educational Researcher*

David M. Quinn
Rossier School of Education
University of Southern California
Waite Phillips Hall – WPH 904E
3470 Trousdale Parkway
Los Angeles, CA 90089

quinnd@usc.edu

Acknowledgments: Funding was provided by a USC Rossier internal research award. Drishti Saxena provided indispensable research assistance. I am grateful to Estela Bensimon, Erika Patall, and Cheryl Ching for feedback on earlier drafts of this article.

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

Abstract

The “achievement gap” has long dominated mainstream conversations about race and education. Some scholars warn that the discourse around racial gaps perpetuates stereotypes and promotes the adoption of deficit-based explanations that fail to appreciate the role of structural inequities. I investigate through three randomized experiments. Results indicate that a TV news story about racial achievement gaps (versus a control or counter-stereotypical video) led viewers to express more exaggerated stereotypes of Black Americans as lacking education (study 1: $ES=.30$ SD; study 2: $ES=.38$ SD) and may have increased viewers’ implicit stereotyping of Black students as less competent than White students (study 1: $ES=.22$ SD; study 2: $ES=.12$ SD, *n.s.*). The video did not affect viewers’ explicit competence-related racial stereotyping, the explanations they gave for achievement inequalities, or their prioritization of ending achievement inequalities. After two weeks, the effect on stereotype exaggeration faded. Future research should probe how we can most productively frame educational inequality by race.

Key words: “stereotyping,” “implicit stereotypes,” “academic expectations,” “achievement gap,” “opportunity gap,” “educational equity”

Experimental Effects of “Achievement Gap” News Reporting on Viewers’ Racial Stereotypes, Inequality Explanations, and Inequality Prioritization

News consumers often encounter headlines such as “Achievement Gap between White and Black Students Still Gaping” (Camera, 2016) or “Education Racial Gap Wide as ever According to NAEP” (Lee, 2014). This “achievement gap discourse”¹ (AGD) (Carey, 2014) has long dominated mainstream conversations about race and education in the US (Ladson-Billings, 2006). While these academic inequalities must be known if they are to be redressed (Harper, 2015; Perry, 2003), a growing number of commentators warn that the AGD’s framing of racial inequality may have unintended consequences (e.g., Ladson-Billings, 2006). By framing the issue as one of individual achievement, the AGD may pathologize students while shifting attention away from the structural inequities that produce unequal opportunities (Ladson-Billings, 2006). Relatedly, the AGD feeds into stereotypes of Black, Latinx, and Native American students as less intellectually capable than White and Asian students (e.g., Cross, 2007; Harper, 2015; Perry, 2003). Among other negative consequences, this sets the conditions for stereotype threat (Steele, 2011).

While theory suggests the AGD may produce these adverse effects (e.g., Carey, 2014; Cross, 2007; Gutierrez, 2008; Harper, 2015; Ladson-Billings, 2006; Love, 2004; Milner, 2012; Perry, 2003), experimental evidence is scant. As a field, we must better understand the effects of the AGD in order to conduct a collective conversation that most effectively advances equity and excellence in education. To my knowledge, the present study offers the first experimental investigation into the causal effects of AGD news stories on people’s implicit and explicit stereotypes, issue prioritization, and their explanations for racial inequalities in educational outcomes.

Background

Motivating the “Achievement Gap” Focus

Many advocates – including past US presidents and education secretaries - have framed the “achievement gap” as “the civil rights issue of our time” (e.g., NBC News, 2014).² One reason for this framing is that schools have the potential to either reproduce or interrupt social stratification. For example, data from the early 1990s showed that all of the Black-White wage gap for women, and much of the wage gap for men, could be explained by test scores (Neal & Johnson, 1996). As such, there is hope that closing Black-White test score gaps will make substantial progress towards ending Black-White inequality in economic outcomes. The general sentiment is well-captured by Christopher Jencks’s and Meredith Phillip’s (1998) argument that “reducing the black-white test score gap would probably do more to promote [racial equality] than any other strategy that commands broad political support” (p.3-4).

Given the implications for racial justice, advocates and researchers focus on achievement gaps as key educational outcomes (e.g., Quinn, 2015; Quinn & Le, 2018). The No Child Left Behind (NCLB) Act of 2001 attracted bipartisan support largely due to its requirements for disaggregating test scores by subgroup (including race) and for holding schools accountable to closing gaps (Hess & Petrilli, 2005). After a long history of US schools under-serving students of color and poor students, NCLB’s emphasis on gap-closure was seen by many civil rights groups as an important victory (Darling-Hammond, 2007). Although much of NCLB eventually became broadly unpopular, civil rights groups continued to support testing and disaggregation requirements when the “Every Student Succeeds Act” came up for a vote in 2015, fearing that turning the spotlight off gaps would lead to inequitable practices (Brown, 2015). Publicizing

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

between-group achievement inequalities, then, is often part of a strategy to make educational equity a national priority.

Deficit Orientations and the Achievement Gap Discourse

Other scholars express concern that the achievement gap Discourse assumes a deficit orientation. For example, Harper (2015) argues that the “overwhelming majority of published scholarship on urban high schools in the United States focuses on problems of inadequacy, instability, underperformance, and violence” and that “images of Black and Latino male students in inner-city schools often manufacture dark, hopeless visualizations of imperiled youth and educational environments” (p. 139). Ladson-Billings (2006) argues that the “achievement gap” framing inappropriately focuses attention on student performance when the focus should be on the disparities in per-pupil funding, healthcare, wealth, or other underlying inequities that lead to racial differences in student outcomes (Ladson-Billings, 2006). This misdirected focus may promote explanations in which students and their families – as opposed to systems of power and privilege - are blamed for students’ perceived “under-achievement.” Furthermore, the simple act of comparing Black, Latinx, and Native American students to White students may reinforce an assumption that White achievement is the standard to which other groups should aspire (e.g., Carey, 2014; Cross, 2007; Harper, 2015; Gutierrez, 2008; Ladson-Billings, 2006; Love, 2004).

Consistent with these concerns, my coauthors and I found in a recent study that the term “achievement gap” itself may be detrimental (Quinn, Desruisseaux, & Nkansah-Amankra, 2019). In a randomized survey experiment with a national sample of teachers (n=1,549), we found that respondents gave lower priority ratings to “closing the racial achievement gap” compared to the conceptually synonymous “ending racial inequality in educational outcomes.” This result is consistent with the notion that deficit-based framing decreases the perceived importance of these

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

disparities, compared to a framing with social justice connotations. Beyond the effect of this small change of phrase, we know very little about the broader effects of more fully-developed achievement gap Discourses.

Perpetuating explicit and implicit stereotypes. One way in which the AGD may do harm is by perpetuating explicit or implicit stereotypes of students of color (Perry, 2003). Holding an explicit stereotype means that one consciously endorses the stereotype, while an implicit stereotype is one that is not identifiable through introspection (Greenwald & Banaji, 1995). Implicit stereotypes can be automatically activated in one's mind, leading to biased behaviors or judgments (Devine, 1989). Thus, people can exhibit "implicit bias" when they do not consciously endorse the stereotype from which it stems (Devine, 1989). In "dual process" models in psychology (Payne & Bishara, 2009), implicit cognition is more likely to drive one's behavior when one is unable or uninclined to think carefully, such as when under time constraints, stress, or when fatigued. Whether explicit or implicit, stereotypes put a "threat in the air" that negatively impacts student performance (Steele, 2011).

A complete theory of stereotypes must answer questions such as why some dimensions of social identity (but not others) are subject to stereotypes, what determines stereotype content, and what role stereotypes play in social reproduction (Bigler & Liben, 2006). However, when it comes to the narrower question of how the AGD may affect the public's stereotypes, the associative learning model and the "representativeness heuristic" (Bordalo, Coffman, Gennaioli, & Shleifer, 2016; Kahneman & Tversky, 1972) are useful.

Associative learning. The associative learning model holds that people learn through association. Learned associations may be between a stimulus and a positive or negative emotional valence, or between a stimulus and an idea (Le Pelley, Reimers, Calvini, Spears,

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

Beesly, & Murphy, 2010). With racial stereotypes, for example, frequent exposure to the association of Black students with low educational achievement may lead one to develop the explicit or implicit stereotype that Black students are unintelligent (an idea), along with an accompanying negative valence. Associations can be learned through exposure to pairings of a neutral stimulus with a positively- or negatively-valenced stimulus (Olson & Fazio, 2001) (e.g., Black students depicted along with signifiers of low performance), or by repeated evaluative statements (Kurdi & Banaji, 2017) (e.g., direct statements about Black students being low-performing).

Associative learning techniques may also help people unlearn stereotypes they have been conditioned to hold. For example, in a study of 17 interventions aimed at reducing implicit bias, Lai and colleagues (2014) found that eight were effective. The most effective were interventions that employed associative techniques such as exposing participants to counter-stereotypical examples. As such, counter-stereotypical depictions of Black students might help counteract any AGD effects.

Stereotypes and the “representativeness heuristic.” According to the representativeness heuristic (Bordalo et al., 2016; Kahneman & Tversky, 1972), stereotypes often emerge from between-group comparisons, and may possess a “kernel of truth” (Bordalo et al., 2016). Yet stereotypes do not necessarily form around the traits that are most common within a group; rather, stereotypes often form around the traits on which groups differ (Bordalo et al., 2016). For example, Florida residents are stereotyped as elderly not because most Floridians are elderly (only 17.4% are 65+), but because a higher percentage of Floridians are elderly compared to the US as a whole (13.1%; Bordalo et al., 2016). Such distinguishing features are over-weighted in

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

people's minds, leading to an exaggerated perception of the magnitude of between-group differences, which may in fact be quite small (Bordalo et al., 2016).

The representativeness heuristic predicts that the between-group comparisons characteristic of the AGD will lead people to develop exaggerated stereotypes of Black Americans as lacking formal education. When this stereotype distortion occurs in the absence of an understanding of structural inequalities, there is also danger that these exaggerated stereotypes will be accompanied by an increase in the extent to which people perceive Black Americans as less intelligent or less hardworking than White Americans.

Summary

Achievement gap statistics are closely tracked and reported on, under the assumption that such monitoring is helpful for efforts to end educational inequality. However, some scholars warn that the way in which the AGD frames racial inequality may have unintended negative consequences. The AGD may perpetuate implicit and explicit stereotypes of racially minoritized students as academically incapable, and lead to exaggerated perceptions of minoritized Americans as uneducated. There is also concern that the AGD's focus on student outcomes rather than structural inequities may encourage explanations that blame students and their families. However, we lack experimental evidence on the AGD's effects. In this article, I report results from three experiments bringing evidence to bear on these questions.

Study 1

In Study 1, I ask:

RQ1: Does exposure to a TV news story about racial test score gaps, versus a counter-stereotypical video, affect viewers' implicit stereotypes, perceived sense of stereotype representativeness, prioritization of racial achievement inequality, or explanations for Black-White achievement inequality?

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

Methods

Participants. Study 1 participants ($n=565$)³ were recruited for this web-based experiment through Qualtrics survey panels, limited to US respondents age 18 or over. Research has shown these panels to be relatively representative of the US population demographically and politically (Boas, Christenson, & Glick, 2018); however, my analytic sample had a smaller share of White (~65%) and larger share of female (~68%) respondents compared to the 2010 U.S. Census (72% and 50.8%, respectively). The modal age category in the sample was 30-39 (~43% of sample), and approximately 11% reported working in education (see Table 1).

Videos. Respondents were randomly assigned to view either an AGD video or a counter-stereotypical video online. Respondents entered key words describing their video as an attention/comprehension check after viewing.

The AGD video was a 129-second clip of a TV newscast from a CBS affiliate in Minnesota (WCCO - CBS Minnesota, 2016). The newscast began with two anchors at the desk, one of whom opens with the narration, “Disappointing numbers out today show the wide achievement gap in Minnesota between White and minority students is not getting any smaller.” The clip continues with narration, images of students in school, and interviews with teachers and school officials. Other key pieces of narration include statements such as, “...most disappointing is the failure to close the achievement gap. The goal was to trim it in half by next year – that’s now a long shot. The gap between White and Black students is essentially unchanged. Seventy percent of White kids are proficient in both subjects compared to thirty-two percent of all Black children.” The between-group test score comparison and the bleak language in this clip are typical features of the AGD (Carey, 2014; Gutierrez, 2008; Harper, 2015).

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

The counter-stereotypical video was a 213-second promotional video from the Promise Academy of Harlem Children's Zone. The video was counter-stereotypical in its anti-deficit presentations of Black students as studious, academically ambitious, and engaged with their vibrant and positive school environment. The video featured clips of students of various ages, in school wearing school uniforms, discussing their academic goals and sharing what they like about their school (Harlem Children's Zone, 2015). See online Appendix A for video transcripts.

This design does not allow for isolating the effects of specific aspects of the AGD or CS video. As such, the estimated AGD effects encompass the collection of reported facts, as well as the tone and sentiment of the discussion.

Measures.

Bias in perceived stereotype representativeness. As discussed above, the AGD may cause exaggerated perceptions of the extent to which racial stereotypes regarding academics are representative of minoritized Americans. To measure this, I administered the following item: "The national high school graduation rate for White students is 86%. What is your best guess of what the national high school graduation rate is for Black students? Type the percentage in the box below." A respondent's guess can be interpreted roughly as their estimate of the probability that a randomly chosen Black adult American holds a high school degree (a proxy for how educated respondents think Black Americans are). Consequently, the larger the respondent believed the Black-White difference to be compared to the actual difference, the more bias in the extent to which the respondent stereotypes Black Americans as lacking education (compared to the actual Black graduation rate of 78% [Murnane, 2013])⁴. An important feature of this measure is that it registers whether respondents (over-)generalize from information about the

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

performance of children on standardized tests in one US state to the graduation rates of adults nationally. Later, I consider advantages and limitations of this measure.

Implicit stereotypes. I adapted the traditional implicit association test (IAT) to measure respondents' automatic associations between race (Black/White) and competence (using the iatgen online software; Carpenter et al., 2018). The traditional Black/White IAT is weakly predictive of various biased behaviors at the individual level (Greenwald, Poehlman, Uhlmann, & Banaji, 2009). In recent theory, the IAT is not conceived of as measuring fixed attitudes of individuals. Rather, implicit bias is seen as “a social phenomenon that passes through the minds of individuals” which “exists with greater stability in the situations they inhabit” (Payne, Vuletich, & Lundberg, 2017, p.5). This explains why individual-level IAT scores are relatively weak predictors of individuals' behaviors, compared to the magnitude of the associations between group-level IAT averages (such as organization-, county-, or country-level scores) and group-level racial inequalities (Jost, 2019; Payne et al., 2017). As such, the motivation for examining the effect of the AGD on IAT scores stems less from the possibility that exposure to a single AGD story may have long-lasting effects on individuals' implicit racial attitudes. Rather, AGDs may lead to higher levels of bias in contexts in which they are encountered.

On my adapted IAT, positive “*d*-scores” indicate pro-White bias (i.e., implicit stereotypes of White students as more competent and Black students as less competent), negative scores represent pro-Black bias, and zero represents neutrality (split-half reliability = .849, error rate = .084). For my analyses, I divide *d*-scores by the overall sample SD. See Appendix B for detail on this IAT and its development.

Achievement inequality prioritization and explanations. For these outcomes, I adapted measures from Valant and Newark (2016). To measure issue prioritization, respondents rated

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

the extent to which they believed it was a national priority to close the Black/White achievement gap (1=not a priority to 5=essential). Although using this item allows me to build on past research, a drawback is that the “achievement gap” term, as noted, has been shown to elicit lower priority levels than the term “inequality in educational outcomes” (Quinn et al., 2019). In the present study, this could pose a threat to the inference of interest if the item wording were to interact with treatment condition.

To express their explanations for achievement inequalities, respondents rated the extent to which they believed the following factors were responsible for Black/White achievement gaps (5-point scale, with order randomized): School quality, student motivation, parenting, discrimination and racism, genetics, neighborhood environments, home environments, and income levels. I used PCA to create an index in which higher scores represent stronger endorsement of non-structural explanations (genetics, motivation, parenting, and home environments; see Appendix C for detail).

Analysis. For each outcome, I test the causal effect of being randomly assigned to the achievement gap (AG) video versus the counter-stereotypical (CS) video by fitting OLS regression models. The key predictor is a binary indicator for random assignment to the AG group; I include a set of demographic controls to improve the precision of estimated treatment effects.

Results

Descriptive statistics. In Table 1, I present descriptive statistics by treatment condition for sample demographics and outcome measures. As expected given the randomization, groups were demographically similar. Immediately, we learn that both groups, on average, showed

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

implicit stereotypes of Black students as less competent than White students (see positively-signed IAT *d*-scores) and downward bias in their graduation rate guesses for Black students.

<Table 1>

Implicit stereotypes. In Table 2, I present treatment effect estimates on the competence IAT⁵ (see Table 1 for unadjusted t-tests for each outcome). In column 1, the AG coefficient indicates that random assignment to the AGD video increased respondents' implicit stereotyping by .22 *SD*.

<Table 2>

In exploratory analyses, treatment effects did not vary significantly by demographic subgroup (not shown); however, power to detect interactions may be insufficient. See Appendix D for results broken down by subgroup.

Bias in perceived stereotype representativeness. In column 2 of Table 2, I present AGD effects on respondents' graduation rate guesses for Black Americans. The intercept indicates that White males (indeed, the full sample, as seen in Table 1) in the CS condition dramatically under-estimated the graduation rate for Black students, with an average guess of 55% (actual rate is 78% [Murnane, 2013]). As hypothesized, random assignment to the AGD video decreased respondents' graduation rate guesses (i.e., further exaggerated the stereotype) by 6.69 percentage points ($ES=.30$), or a 30% increase in the average bias.

Exploratory analyses showed no significant effect variation by respondent subgroup (see Appendix D).

Gap prioritization and explanations. In columns 3-4 of Table 2, I present AGD effect estimates on gap prioritization and the “non-structural” explanation index (see Appendix C for results with individual gap explanation items; no effect was significant at $p<.05$). I find no

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

evidence that being randomly assigned to watch the AGD video (versus the CS video) affected these outcomes.

Study 2

The observed effects on implicit stereotypes and perceived stereotype representativeness raise the question of whether these effects were driven by the AGD video increasing biases or the CS video decreasing biases. In Study 2, I explored this question with a new sample. I also replicate Study 1 effects and test whether effects faded after two weeks. I ask:

RQ2a: Compared to a control video, does exposure to an AG news story or CS video affect viewers' implicit stereotypes or perceptions of stereotype representativeness?

RQ2b: If so, are effects maintained after two weeks?

Procedures

I recruited U.S. respondents through Amazon Mturk and paid them \$1.50 for participating. Mturk samples have been found to be more representative than in-person convenience samples (Berinsky, Huber, & Lenz, 2012).

Each respondent ($n=723$)⁶ was randomly assigned to one of three videos: the AG or CS video from Study 1, or a third control video. The control was a 175-second Khan Academy (KA) video describing an application of the Pythagorean Theorem (Khan Academy, 2011), chosen because it connects to the field of education but makes no allusions to race or achievement (and is of similar length to the other videos). Respondents completed the key word attention/comprehension check and the graduation rate guess item from Study 1, followed by the competence IAT.

Demographic information was not collected during the first round of Study 2, but was collected during the follow-up (33% response rate). The characteristics of this sample differed in

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

some respects from the Qualtrics sample: Overall, the MTurk follow-up respondents were 7.6% Asian American, 3.8% Black, 3% Latinx, 78.5% White, 40% female, and the modal age group (at 36%) was 30-39. Experimental conditions did not differ on any of these characteristics (though the extent to which this represents the initial sample is unknown; see Appendix E for full descriptive statistics).

Results

Implicit stereotypes. In Column 1 of Table 3, I present the regression results for the competence IAT (see endnote 6 for detail on incalculable IAT scores; rates of invalid scores did not differ across conditions). Here, the KA video is the reference group (with CS and AG estimates shown in relation to the KA video). On average, control group participants showed significant implicit stereotyping of Black students as less competent than White students (intercept = .43). As indicated by the model F-statistic ($F=0.67$), condition had no significant effect on competence IAT scores, and the coefficients for AG and CS demonstrate that neither group's IAT scores differed significantly from the control group's.

<Table 3>

In Columns 3 and 4 of Table 3, I estimate effects on the IAT while pooling the Study 1 sample with the AG and CS groups from Study 2 (KA condition excluded; I control for sample with an indicator, and CS is the reference group). After pooling samples, the significant effect observed in Study 1 on IAT scores is attenuated, but remains statistically significant at .135 *SD*. As seen by the interaction term included in Column 4, the effect of the AGD video did not differ significantly across samples.

Bias in perceived stereotype representativeness. Column 2 of Table 3 shows the regression results for the graduation rate guess. Condition had a significant effect on the

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

magnitude of respondents' racially biased academic stereotypes ($F=14$). Participants in the control group (KA video) under-estimated the Black graduation rate, with an average guess of 62%. The coefficient for AG shows that being randomly assigned to the news clip decreased respondents' graduation rate guesses for Black students (i.e., increased bias) by 7 percentage points compared to the control video ($ES=-.38$), or a bias increase of 44%. However, random assignment to the CS video had no significant effect on bias. In other words, the AGD seems to have increased bias but counter-stereotypical media did not decrease bias. As seen in Column 6 of Table 3, effects did not differ significantly across samples.

Follow-up analyses. In order to determine whether the AG effects on perceived stereotype representativeness were sustained over time, I sent a follow-up invitation to Study 2 Mturk respondents two weeks after they completed the original survey. Respondents were paid 12 cents to answer the same graduation rate item, followed by three demographic questions (race/ethnicity, gender, age range). Two hundred thirty-seven respondents participated at follow-up (33% response rate overall; 31%, 31%, and 36% for the KA, CS, and AGD videos respectively).

Effects were not sustained at follow-up. This appears to be due to respondents' guesses returning to baseline after having been influenced downward by the AGD video, rather than to differential selection into the follow-up sample (See Appendix E).

Study 3

Studies 1 and 2 show robust evidence that the AGD video impacted (in the short-term) the extent to which respondents perceived the "high school drop-out" stereotype as representative of Black Americans, and (weaker) evidence that the video affected implicit competence-related stereotypes. In Study 3, I tested whether the video affected respondents'

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

general explicit racial stereotypes regarding intelligence and competence. I also sought to replicate the Study 1 prioritization result, now with an index formed by averaging five items (Cronbach's $\alpha = .95$; see Appendix F).

For Study 3, I recruited a new MTurk sample of $n=600$ (yielding power of .80 to detect an effect of .23 SD; see Appendix F for descriptive statistics, balance checks, and detail). Respondents were randomly assigned to view either the AGD video or Khan Academy control video, then completed the key word comprehension check, followed by eight stereotype items, five priority items (with order randomized), and demographic items. The stereotype items (adapted from the General Social Survey) asked respondents to rate Black and White Americans on 7-point bipolar scales for the following constructs: intelligent/unintelligent, hard-working/lazy, competent/incompetent, capable/incapable (where higher ratings represented the more positive pole). I created two composites with these items: 1) respondents' mean ratings of Black Americans on the four items (Cronbach's $\alpha = .91$), and 2) respondents' mean White-Black rating difference on the items (Cronbach's $\alpha = .89$; see Appendix F).

In Table 4, I present effect estimates for the AGD video on respondents' explicit competence stereotypes. In both conditions, respondents rated Black Americans lower than White Americans, but neither stereotype measure was significantly affected by the AGD video. Replicating Study 1, effects on the prioritization index were also null.

<Table 4>

Discussion

I report results from three experiments on the effects of the achievement gap Discourse (AGD). After being randomly assigned to view an AGD TV news report about test score gaps, participants expressed larger downward bias in their perceptions of Black Americans' education

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

levels (Studies 1 and 2), and may have increased their implicit stereotyping of Black students as less competent than White students (Studies 1 and 2). However, I did not find evidence that the AGD affected viewers' explicit competence-related racial stereotyping (Study 3), the explanations they offered for racial inequality in educational outcomes (Study 1), or the extent to which they prioritized racial achievement inequalities (Studies 1 and 3). Two weeks after watching the AGD video, the additional downward bias in viewers' perceptions of Black Americans' education levels had returned to baseline bias levels among a subsample of follow-up respondents (Study 2). Unlike the AGD video, a counter-stereotypical video did not affect viewers' implicit stereotypes or perceived stereotype representativeness (Study 2).

Implications

The findings for viewers' perceptions of Black Americans' education levels are consistent with the representativeness heuristic. As described earlier, this theory holds that stereotypes often emerge when groups differ on some trait. People tend to develop an exaggerated perception of how representative the trait is of the stereotyped group. On average, respondents exaggerated the representativeness of the "high school drop-out" stereotype for Black Americans. Importantly, results suggest the AGD may contribute to this exaggerated stereotype: After exposure to the AGD, participants exhibited an increased downward bias in their estimates of the probability that a random Black American would be a high school graduate. For Black Americans, this means being subjected to a stronger stereotype.

Economic models of statistical discrimination suggest potential implications of this finding. Statistical discrimination theory poses that people hold priors regarding how various characteristics are distributed within and between racial groups (e.g., Guryan & Charles, 2013). People informally apply these priors when making inferences about individuals, leading to

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

discriminatory behaviors or judgments (Guryan & Charles, 2013). The present study suggests that when it comes to the distribution of education credentials by race, people's priors are biased downward for Black Americans, and even more strongly biased downward after exposure to the AGD. With larger bias in people's priors, we are likely to find more instances of racial discrimination. This could conceivably play out in various settings, from employers' decisions about who to target for recruitment, to workplace microaggressions in which Black professionals are mistakenly assumed to hold lower-status jobs.

While the AGD may increase bias in people's perceived stereotype representativeness, individuals will nevertheless vary in their explanations for educational attainment differences by race, and in their levels of concern over the disparity. One fear has been that the AGD may lead people to infer racial differences in intelligence, making them less inclined to prioritize ending achievement inequalities. However, others might infer structural explanations after exposure to the AGD and become more motivated to combat inequities. In the latter case, the AGD would have beneficial effects despite the magnified stereotypes. Yet I find no evidence for either of these scenarios. In Study 1, the AGD video did not significantly affect viewers' explanations for existing outcome disparities, be those explanations genetics, individual motivation, racism, school quality, or others. Similarly, the video did not affect explicit racial stereotypes regarding intelligence or competence (Study 3), suggesting that the effect on viewers' graduation rate guesses was not due to a change in viewers' beliefs about Black Americans' capabilities. Rather, the AGD only impacted participants' perceptions of the magnitude of the graduation rate differential - it did not impact the extent to which participants believed any specific factor was responsible.

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

Given that AGDs may, in theory, cause people to de-prioritize achievement inequality, it is somewhat reassuring that the null effect on the prioritization item replicated across two samples (though confidence intervals are wide). At the same time, it is disappointing that informing participants about educational outcome inequalities did not lead them to place a higher priority on these inequalities. However, it is possible that drawing attention to the broader injustices that contribute to unequal outcomes would galvanize more support for equity-advancing efforts. As alluded to above, it is also possible that a differently-worded survey item would have registered an effect. In particular, by asking respondents to rank the priority of “closing the Black/White achievement gap” (as opposed to, e.g., “ending Black/White inequality in educational outcomes”), the item wording may have counteracted any negative effect of the AGD video by depressing priority levels in the control group.

Given the imperative of equalizing educational opportunity by race, the effects observed here on exaggerated stereotype representativeness and (potentially) on implicit stereotypes should lead us to consider carefully how we frame educational inequalities. These findings do not mean that we should cease all measuring or reporting on between-group differences in outcomes. Between-group comparisons are inherent to questions of racial equity, and therefore ceasing all between-group comparisons would mean ignoring questions of equity. Rather, what we need is a better understanding of how certain ways of framing inequalities may be more or less impactful on people’s racial attitudes or stereotypes, and whether, and in what circumstances, any potential negative effects may be outweighed by potential positive effects. With more research, we may uncover better ways of advancing equitable policy without also perpetuating harmful stereotypes.

Limitations and Future Research

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

One limitation of this work relates to the nature of the samples, which are not designed to be nationally-representative. I cannot rule out the possibility that effects are idiosyncratic to the types of respondents available through Qualtrics panels or Mturk. We cannot know whether the inconsistency across samples on the IAT result was due to true heterogeneity in treatment effects (which I did not find, but for which I may be under-powered), simple chance, or differences in the attentiveness of Qualtrics and Mturk respondents.

Similar to past research (Bordalo et al., 2016), I use a single item to measure stereotype representativeness. An advantage of the single “high school graduation rate guess” item is that it provides a direct, objective, and easily interpretable measure of the high school drop-out stereotype. A disadvantage of the single-item approach is that we cannot know the extent to which the AGD video may impact stereotypes on other indicators of educational attainment.

Conceptual replications will help to further develop our understanding of AGD effects. We learned from this study about the effects of a small set of specific video clips, but our interest is in the broader population of representations across media. Do other similar videos have effects? What are the core mechanisms driving the results, and how can we differentiate between representations that do and do not entrench biases? To what extent might incidental aspects such as tone or presentational style influence AGD effects? Are similar effects to be found in print or audio media? Do effects differ when AGDs are presented through narratives or anecdotes versus through quantitative data? Relatedly, this article only tested effects on Black/White bias. Future work should replicate these results among other racial/ethnic groups and along other social dimensions such as class and gender.

Given that people’s beliefs and values affect the way they respond to information (Chong & Druckman, 2007), AGD effects will likely vary across individuals. As alluded to earlier,

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

people who value racial justice may experience AGD effects on their perceptions of the representativeness of racial stereotypes while also increasing their motivation to end educational inequity. Future research should investigate the ways in which AGD effects may be moderated by individuals' baseline knowledge, beliefs, and attitudes. Relatedly, these baseline factors are likely to influence which media reports individuals consume outside of experimental settings (Peralta, Wojcieszak, Lelkes, & Vrees, 2017). The nature of potential selective exposure, and its role in AGD effects, must also be understood.

Although I did not find that the counter-stereotypical video reduced stereotyping, other examples – or a heavier dose of such examples – may be effective at reducing stereotypes. Finally, research should examine how these processes play out among educators, who are often exposed to achievement gap data. What is the effect of encountering these data in real world settings, and do educators respond to them differently compared to the general public?

Conclusion

Racial inequalities in educational outcomes must be known if they are to be redressed (Harper, 2015). At the same time, the present work lends empirical support to scholars' concerns that achievement gap Discourses perpetuate stereotypes. Through future research, we must deepen our understanding of how various communication frames impact the public's understanding of racial inequities in education, and of how we can most productively conduct a public conversation around equalizing opportunity. Results from such work will have important implications for the way that educational inequalities and opportunity gaps are discussed in policy, practice, research, and media.

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

Notes

¹ I follow Carey (2014), who uses the capitalized “achievement gap Discourse” to refer to Gee’s (1996) notion of Discourse as “language through the lens of social context and broad cultural ideological processes” (Carey, 2014, p. 441).

² While the term “achievement gap” is often used to cover numerous between-group comparisons on various educational outcomes, the language of “civil rights” elicits associations with the Civil Rights Movement, the *Brown* decision, and Black-White inequality generally. Indeed, the search phrase *Black White achievement gap* on Google produces far more results than other comparisons (results returned by search term on 2-2-20: Black White achievement gap: 64m; Hispanic White achievement gap: 5m; Latino White achievement gap: 5.9m; Asian white achievement gap: 21.2m).

³ A target sample of 500 respondents was sought for .80 power to detect an effect of .25 *SD*. Qualtrics, which managed the data collection, obtained a larger sample than anticipated, and ended data collection after 610 respondents submitted surveys (prior to data analysis). Following their protocol, Qualtrics paused data collection after approximately ten percent of the target sample had submitted surveys (in this case, data collection was paused after 45 respondents). At this stage, it became clear that a survey-advance delay had not been programmed into the video page as intended, allowing respondents to advance past the video before its completion. The delay was then added for subsequent data collection. The first 45 responses were discarded, resulting in a final analytic sample of 565. Data checks suggested that the first 45 respondents did indeed skip through the videos: the median total survey time for the first 45 respondents was 105 seconds shorter compared to all subsequent respondents. When the first 45 respondents are included in the analysis, results for the graduation rate guess outcome are nearly identical ($b=-$

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

6.81, $p < .001$) and the effect on the IAT is reduced to 0.142 SD ($p = .075$). To ensure complete data, all survey item responses were forced.

⁴Murnane's (2013) estimates are "adjusted status completion rates" (i.e., excluding GEDs) based on data for 20-24 year-olds from the American Community Survey in 2010. Recent immigrants are excluded from the numbers. See Murnane (2013) for a discussion of the complications involved when estimating graduation rates.

⁵After dropping respondents for whom valid IAT scores could not be calculated (due to excessively rapid responses, ~9% of the sample), the analytic sample for the IAT analyses came to 514 (rates of invalid scores did not differ by condition; females were slightly less likely to have invalid IAT scores than males, but no demographic characteristics interacted with treatment condition when predicting invalid score status).

⁶An initial target sample of $n=600$ was sought (200 per condition). When this quota was reached in MTurk, only 463 of the submissions produced valid IAT scores (others incalculable due to excessive speed); 578 included graduation rate guess responses. In an attempt to reach the target sample size for the IAT analyses, I sought an additional 150 respondents. This yielded 120 additional valid IAT scores and an additional 145 graduation rate guess submissions, resulting in the final analytic samples shown here. For both rounds of data, effects on the graduation rate guesses were negative and significant ($b = -6.61$, $p < .001$ and $b = -8.57$, $p = .021$, respectively) and effects on IAT are positive and non-significant ($b = .07$, $p = .52$ and $b = .28$, $p = .19$, respectively).

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

References

- Berinsky, A.J., Huber, G.A., & Lenz, G.S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis*, 20(3): 351-368.
- Bigler, R. S., & Liben, L. S. (2006). A developmental intergroup theory of social stereotypes and prejudice. In *Advances in child development and behavior* (Vol. 34, pp. 39-89).
- Boas, T.C., Christenson, D.P., & Glick, D.M. (2018). Recruiting large online samples in the United States and India: Facebook, Mechanical Turk, and Qualtrics. *Political Science Research and Methods*, X, 1-19, doi:10.1017/psrm.2018.28
- Bordalo, P., Coffman, K., Gennaioli, N., & Shleifer, A. (2016). Stereotypes. *Quarterly Journal of Economics*, 131, 1753-1794.
- Brown, E. (January 12, 2015). Civil rights groups back standardized testing in No Child Left Behind rewrite. *Washington Post*. Retrieved from: https://www.washingtonpost.com/local/education/civil-rights-groups-back-standardized-testing-in-no-child-left-behind-rewrite/2015/01/12/511c99e4-99fd-11e4-a7ee-526210d665b4_story.html?utm_term=.861b06b08da4
- Chong, D., & Druckman, J.N. (2007). Framing theory. *Annual Review of Political Science*, 10, 103-126.
- Camera, L. (2016). Achievement gap between Black and White students still gaping. *US News and World Report*. Retrieved from <https://www.usnews.com/news/blogs/data-mine/2016/01/13/achievement-gap-between-white-and-black-students-still-gaping>
- Carey, R. L. (2014). A cultural analysis of the achievement gap Discourse: Challenging the language and labels used in the work of school reform. *Urban Education*, 49, 440-468.
- Carpenter, T., Pogacar, R., Pullig, C., Kouril, M., LaBouff, J., Aguilar, S. J., ... Chakroff, A. (2018, April 3). Conducting IAT Research within Online Surveys: A Procedure, Validation, and Open Source Tool. <https://doi.org/10.31234/osf.io/6xdyj>
- Cross, B.E. (2007). Urban school achievement gap as a metaphor to conceal U.S. apartheid education. *Theory Into Practice*, 46(3), 247-255.
- Darling-Hammond, L. (May 2, 2007). Evaluating “No Child Left Behind”: The problems and promises of Bush’s education policy. *The Nation*. Retrieved from: <https://www.thenation.com/article/evaluating-no-child-left-behind/>
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of personality and social psychology*, 56(1), 5.
- 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.
- Greenwald, A.G., & Banaji, M.R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102, 4-27.
- Guryan, J., & Charles, K.K. (2013). Taste-based or statistical discrimination: The economics of discrimination returns to its roots. *The Economic Journal*, 123, F417-F431.
- Gutiérrez, R. (2008). A "gap-gazing" fetish in mathematics education? Problematizing research on the achievement gap. *Journal for Research in Mathematics Education*, 357-364.
- Harlem Children’s Zone. (2015, April 3). *A look inside Promise Academy*. [Video file]. Retrieved from: <https://www.youtube.com/watch?v=e4H-Esejngg&feature=youtu.be>
- Harper, S. R. (2015). Success in these schools? Visual counternarratives of young men of color and urban high schools they attend. *Urban Education*, 50(2), 139-169.

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

- Hess, F.M., & Petrilli, M.J. (2005). The politics of No Child Left Behind: Will the coalition hold? *Journal of Education*, 185(3), 13-25.
- Jencks, C., & Phillips, M. (1998). The Black-White test score gap: An introduction. *The Black-White test score gap*, Ed. C. Jencks & M. Phillips, p1-51. Washington, D.C.: The Brookings Institution.
- Jost, J. T. (2019). The IAT is dead, long live the IAT: Context-sensitive measures of implicit attitudes are indispensable to social and political psychology. *Current Directions in Psychological Science*, 28(1), 10-19.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3, 430-454.
- Khan Academy (2011, November 9). *Pythagorean theorem application*. [Video file]. Retrieved from: <https://www.youtube.com/watch?v=l43fgfJ376c&feature=youtu.be>
- Kurdi, B., & Banaji, M. R. (2017). Repeated evaluative pairings and evaluative statements: How effectively do they shift implicit attitudes?. *Journal of Experimental Psychology: General*, 146(2), 194.
- Ladson-Billings, G. (2006). From the achievement gap to the education debt: Understanding achievement in U. S. schools. *Educational Researcher*, 35, 3-12
- Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J. L., Joy-Gaba, J. A., Ho, A.K., Teachman, B. A., Wojcik, S. P., Koleva, S. P., Frazier, R. S., Heiphetz, L., Chen, E., Turner, R. N., Haidt, J., Kesebir, S., Hawkins, C. B., Schaefer, H. S., Rubichi, S., Sartori, G., Dial, C. M., Sriram, N., Banaji, M. R., & Nosek, B. A. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, 143, 1765-1785.
- Le Pelley, M.E., Reimers, S.J., Calvini, G., Spears, R. Beesley, T., & Murphy, R.A. (2010). Stereotype formation: Biased by association. *Journal of Experimental Psychology: General*, Vol 139(1), 138-161.
- Lee, T. (2014). Education racial gap as wide as ever according to NAEP. MSNBC.com Retrieved from <http://www.msnbc.com/msnbc/student-proficiency-stagnant-race-gap-wide>
- Love, B.J. (2004). *Brown Plus 50 Counter-Storytelling: A Critical Race Theory Analysis of the “Majoritarian Achievement Gap” Story*. *Equity & Excellence in Education*, 37, 227-246.
- Milner, H.R. (2012). Beyond a test score: Explaining opportunity gaps in educational practice. *Journal of Black Studies* 43(6), 693-718.
- Murnane, R.J. (2013). U.S. high school graduation rates: Patterns and explanations. *Journal of Economic Literature*, 51, 370-422.
- Neal, D.A., & Johnson, W.R. (1996). The role of premarket factors in black-white wage differentials. *Journal of Political Economy*, 104, 869-895.
- NBC News. (April 10, 2014). Bush: Achievement gap an urgent civil rights issue. Retrieved from: <https://www.nbcnews.com/politics/politics-news/bush-achievement-gap-urgent-civil-rights-issue-n77416>
- Olson, M. A., & Fazio, R. H. (2001). Implicit attitude formation through classical conditioning. *Psychological Science*, 12, 413–417. <http://dx.doi.org/10.1111/1467-9280.00376>
- Payne, B.K., & Bishara, A.J. (2009). An integrative review of process dissociation and related models in social cognition. *European Review of Social Psychology*, 20, 272-312.
- Payne, B.K., Vuletic, H.A., & Lundberg, K.B. (2017). The bias of crowds: How implicit bias bridges personal and systemic prejudice. *Psychological Inquiry*, 28, 233-248.

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

- Peralta, C.B., Wojcieszak, M., Lelkes, Y., & Vrees, C. (2017). Selective exposure to balanced content and evidence type: The case of issue and non-issue publics about climate change and health care. *Journalism & Mass Communication Quarterly*, *94*, 833-861.
- Perry, T. (2003). Up from the parched earth: Toward a theory of African-American achievement. In Steele, C., & Hilliard, A.G. (2003). *Young, Gifted, and Black: Promoting High Achievement Among African-American Students*. Boston: Beacon Press.
- Quinn, D.M. (2015). Kindergarten black-white test score gaps: Re-examining the roles of socioeconomic status and school quality with new data. *Sociology of Education*, *88*, 120-139.
- Quinn, D.M., & Le, Q.T. (2018). Are we trending to more or less between-group achievement inequality over the school year and summer? Comparing across ECLS-K Cohorts. *AERA Open*, *4*(4), 1-19. DOI: 10.1177/2332858418819995
- Quinn, D.M., Desruisseaux, T.K., & Nkansah-Amankra, A. (2019). Achievement gap language affects teachers' issue prioritization. *Educational Researcher*, *48*, 484-487.
- Steele, C.M. (2011). *Whistling Vivaldi: How stereotypes affect us and what we can do*. W.W. Norton & Company.
- Valant, J., & Newark, D. A. (2016). The politics of achievement gaps: US public opinion on race-based and wealth-based differences in test scores. *Educational Researcher*, *45*(6), 331-346.
- WCCO - CBS Minnesota (2016, July 28). *Testing shows no progress in closing achievement gaps*. [Video file]. Retrieved from:
<https://www.youtube.com/watch?v=Nx4QCh02rCI&feature=youtu.be>

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

Table 1.
Descriptive Statistics by Condition (Study 1).

	Achievement Gap Discourse Video			Counter-stereotypical Video			p
	Mean	SD	N	Mean	SD	N	
<i>Sample</i>							
American Indian	0.010		287	0.004		278	0.332
Asian American	0.049		287	0.072		278	0.248
Black	0.167		287	0.147		278	0.520
Latinx	0.056		287	0.047		278	0.629
Multi-racial	0.049		287	0.090		278	0.054
Other Race	0.003		287	0.011		278	0.301
Pacific Islander	0.000		287	0.000		278	N/A
White	0.666		287	0.629		278	0.371
Female	0.666		287	0.691		278	0.523
Male	0.317		287	0.309		278	0.844
Non-binary	0.017		287	0.000		278	0.027
Age 0-18	0.003		287	0.000		278	0.325
Age 19-29	0.317		287	0.317		278	0.989
Age 30-39	0.425		287	0.442		278	0.678
Age 40-49	0.247		287	0.234		278	0.707
Age 50-59	0.007		287	0.004		278	0.582
Age 60-69	0.000		287	0.004		278	0.310
Educator	0.108		287	0.115		278	0.789
<i>Outcomes</i>							
<i>Stereotypes</i>							
IAT D-score	0.249	0.459	256	0.159	0.434	258	0.023
Grad rate guess	49.369	22.072	287	55.917	21.390	278	0.000
<i>Gap/explanations</i>							
Priority gap	3.836	1.047	287	3.835	1.092	278	0.985
School quality	4.017	0.988	287	3.856	1.068	278	0.063
Motivation	3.763	1.051	287	3.615	1.140	278	0.109
Parenting	4.000	1.034	287	3.849	1.124	278	0.097
Discrimination & racism	3.812	1.137	287	3.712	1.265	278	0.325
Genetics	2.767	1.393	287	2.766	1.479	278	0.998
Neighborhood	4.024	0.951	287	3.975	1.069	278	0.560
Home	4.070	0.955	287	3.924	1.008	278	0.079
Income	3.735	1.128	287	3.809	1.073	278	0.424
Explanation index (std)	0.063	0.963	287	-0.060	1.026	278	0.144
Non-structural explan index (std)	0.029	0.959	287	-0.040	1.040	278	0.412

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

Note. Items without SD are binary indicator variables for row category. *P*-value is for t-test of null hypothesis of equal means across conditions. Priority gap = how much of a priority respondent believes it is to close Black/White achievement gap (1=not a priority, 2=low priority, 3=medium priority; 4=high priority; 5=essential). Gap explanations Qs give extent to which respondent believes that factor is responsible for Black/White achievement gap (1=not at all; 2=slightly; 3=somewhat; 4=quite; 5=extremely).

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

Table 2.

OLS Regression Models Estimating Treatment Effects on Implicit Stereotypes (IAT), Perceived Stereotype Representativeness (Grad Guess), Gap Prioritization, and Gap Explanations (Study 1).

	(1) IAT D-score	(2) Grad Guess	(3) Gap Priority (Std.)	(4) Non-structural Explanations Index (Std.)
AG	0.224** (0.0820)	-6.686*** (1.850)	-0.0149 (0.0844)	0.0747 (0.0838)
Black	-1.043*** (0.113)	-1.268 (2.571)	0.283* (0.117)	-0.414*** (0.116)
Latinx	-0.335~ (0.187)	-1.353 (4.196)	-0.177 (0.192)	0.0852 (0.190)
Asian	-0.172 (0.175)	-4.783 (3.904)	0.0706 (0.178)	-0.132 (0.177)
Other Race	-0.922 (0.652)	-16.89 (10.94)	0.696 (0.499)	0.358 (0.496)
Multi-racial	-0.674*** (0.158)	-2.339 (3.678)	-0.152 (0.168)	-0.392* (0.167)
American Indian	-0.962* (0.462)	10.70 (10.94)	0.236 (0.499)	-0.635 (0.496)
Female	0.0292 (0.0898)	2.900 (1.981)	0.193* (0.0904)	-0.165~ (0.0898)
Non-binary	-0.865* (0.420)	-8.909 (9.929)	0.928* (0.453)	-0.840~ (0.450)
Constant	0.586*** (0.0917)	54.86*** (2.033)	-0.162~ (0.0928)	0.174~ (0.0921)
<i>N</i>	514	565	565	565
<i>R</i> ²	0.176	0.037	0.030	0.045
<i>F</i>	11.99	2.390	1.938	2.907

Note. Standard errors in parentheses. AG= achievement gap video (1=AG, 0=counter-stereotypical video). All predictor variables are binary indicators for the named group. IAT D-scores are divided by the overall sample sd to facilitate interpretation. “Gap priority (Std)” and “Non-structural explanations (Std)” outcomes are standardized to mean=0, SD=1.

~ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

Table 3.

OLS Regression Models Estimating Treatment Effects on Implicit Stereotypes (IAT) and Perceived Stereotype Representativeness (Grad Guess) for Study 2 (Columns 1 & 2) and for Pooled Study 1 and Study 2 Samples (Columns 3-6).

	(1)	(2)	(3)	(4)	(5)	(6)
	Study 2 Sample		Pooled Sample Study 1 & 2			
	IAT D-score	Grad Guess	IAT D-score	IAT D-score	Grad Guess	Grad Guess
AG	0.117 (0.101)	-7.036*** (1.645)	0.135* (0.0668)	0.201* (0.0890)	-7.124*** (1.218)	-6.548*** (1.667)
CS	0.0672 (0.102)	0.747 (1.675)				
Mturk Sample			0.0659 (0.0673)	0.142 (0.0957)	6.541*** (1.220)	7.175*** (1.750)
AG*Mturk				-0.150 (0.135)		-1.236 (2.442)
Constant	0.430*** (0.0735)	62.34*** (1.195)	0.388*** (0.0555)	0.355*** (0.0628)	56.21*** (1.038)	55.92*** (1.188)
<i>N</i>	583	723	913	913	1059	1059
<i>R</i> ²	0.002	0.037	0.006	0.007	0.056	0.056
<i>F</i>	0.672	14.00	2.549	2.116	31.16	20.85

Note. AG, CS=dummy variables for achievement gap and counter-stereotypical videos, respectively. Study 2 sample includes data from three conditions: AG, CS, and the Khan Academy (KA) video (omitted reference group). Pooled sample excludes Khan Academy, which was not a Study 1 condition (CS is omitted reference group for pooled sample). IAT sample sizes by condition for Study 2: KA= 184, CS=195, AG=204; Grad Guess sample sizes by condition for Study 2: KA=229, CS=238, AG=256. D-scores are divided by the overall sample sd. Mturk Sample= indicator variable for whether observation was from Mturk (vs. Qualtrics sample).

~ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

EFFECTS OF ACHIEVEMENT GAP NEWS REPORT

Table 4.

OLS Regression Models Estimating Treatment Effects on Explicit competence-related Racial Stereotypes and Inequality Prioritization (Study 3)

	(1) Black Stereotype	(2) W-B Stereotype Diff.	(3) Priority
AG	-0.0987 (0.0960)	0.0686 (0.0851)	0.143 (0.0920)
Constant	4.906*** (0.0683)	0.362*** (0.0606)	3.555*** (0.0655)
<i>N</i>	600	600	600
<i>R</i> ²	0.002	0.001	0.004

Note. Standard errors in parentheses. AG= achievement gap video (1=AG, 0=Khan Academy video). Black stereotype=mean score on four 7-point bipolar scales rating Black Americans on the constructs: intelligent/unintelligent, hard-working/lazy, competent/incompetent, capable/incapable (higher ratings represent the more positive pole). W-B Stereotype Diff=mean White-Black difference on each stereotype item (such that positive values indicate a pro-White stereotype). Priority= mean score on five items measuring the extent to which respondents believe that closing Black/White achievement gaps is a priority (see Appendix F for detail).

~ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Appendix A. Video Transcripts.

News Clip (Achievement Gap Discourse Video)

Anchor 1, Frank Vascellaro (on screen): Disappointing numbers out today show the wide achievement gap in Minnesota between white and minority students is not getting any smaller.

Anchor 2, Amelia Santaniello (on screen): The Department of Education released results from 2016 student testing and it shows virtually no progress in reading and math scores. It is now the third straight year of stagnation and today Education Commissioner Brenda Cassellius expressed disappointment. And as Bill Hudson explains, it shows that the work of closing the gap extends far beyond any classroom.

Bill (narrating): Student testing has always been a part of education. Knowledge in reading, math, and science reveals a lot about students and their instruction.

Commissioner Cassellius: We're all wondering why can't we get at this disparity.

Bill (narrating): While disappointed in the most recent MCA scores, Minnesota's education commissioner says it also reflects more rigorous testing.

Commissioner: Teachers need better resources and curriculum development and professional development so that they can understand these very complex and different standards than they taught maybe 20 years ago.

Bill (narrating): While reading proficiency scores were up slightly to sixty percent of all students, math proficiency dropped: it fell a point down to 61 percent but most disappointing is the failure to close the achievement gap. The goal was to trim it in half by next year - that's now a long shot. The gap between white and black students is essentially unchanged. Seventy percent of white kids are proficient in both subjects compared to thirty-two percent of all black children.

Mark Westpfahl: What is the actual purpose of this test, what's it measuring?

APPENDICES: EFFECTS OF ACHIEVEMENT GAP NEWS REPORTING

Bill (narrating): Middle school educator Mark Westpfahl believes we're over-testing at the cost of teaching.

Mark: And you're getting in the mode of always getting ready to do tests rather than learning the skills that are necessary to achieve.

Bill (narrating): Still one thing is abundantly clear: schools can't close the gap alone - that will take a wider effort beginning earlier and at home. Bill Hudson WCCO 4 News.

Anchor 2, Amelia Santaniello (on screen): Results showing how well individual schools are performing – the so-called multiple measurements rating – will be released in early September.

Harlem Children's Zone Promise Academy (Counter-stereotypical Video)

~("With my own Two Hands," performed by Ben Harper, plays in background)~

Song: "I can change the world with my own two hands, make a better place with my own two hands, make a kinder place with my with my own two hands."

Student 1: I like this school 'cause it's giving us more knowledge and it's like kind of strict but that's good 'cause it's preparing us for adult world.

Student 2: What I like about school is that my teacher is beautiful, Ms. Hardy and my classmates beautiful.

Student 3: I think this school rocks.

~ (Song continues) ~

Student 4: Everybody tries to help us in the best way they can.

Student 5: If you need extra help, like away from the class, they have that.

Student 6: I like how many classes we have, they made us - they gave us better and longer class periods.

Student 2: Writing, reading and social studies.

Student 7: And there's a lot of networking going on in Promise Academy, so my mind's expanded.

Student 1: African American history is teaching us about our American culture.

Student 6: I wanna be either an actor or a dancer.

Student 2: I wanna be in Jeopardy. You could be smart like that.

Student 5: Teach kids in like 5th-6th grade

Student 7: I wanna be a psychiatrist

Student 2: A famous dancer.

APPENDICES: EFFECTS OF ACHIEVEMENT GAP NEWS REPORTING

Student 1: A paleontologist

Student 8: When I grow up I'm going to be a doctor

Student 3: I wanna be an astronaut when I grown up

Student 2: When I graduate, I'm gonna get married, get my kids and I'mma be set.

Student voiceover: When I first saw this school the first thing I thought that it was very colorful, amazing and bright

Student voiceover: Every floor is a different color

Student 8: It's Yellow, blue and white

Student voiceover: There's just more space, there's more options

Student 2: And when you put it all together - it looks like, like it's a mansion.

~(Song Continues)~

Student 4: I love the dance room, we have mirrors, we have the ballet pole...

Student 2: They have a big park that we play in and it's so fun.

Student 5: Gym - because sometimes we have to go to the patio - and we'll play football, it just gives us a chance to just like let out all our energy.

Student 2: They let you do basketball, football, anything you want but you always gotta listen in gymnastics.

Teacher: Three Two One -

(Student choral chant of school anthem)

Khan Academy Pythagorean Theorem (Control Video, Studies 2 & 3)

Voiceover: A 60-foot ladder is put up against a building. The base of the ladder is eight feet away from the building. How high will the ladder reach?

So let's draw this scenario here - so let's say that this is the ground and this is the building, this is the building. And then we're going to have a 60-foot ladder and it's leaning up against this building.

So the length of this ladder is 60 feet that is a 60-foot ladder and then the base of the ladder is eight feet away from the building. So this distance right over here is eight - is eight feet.

And they say "how high will the ladder reach?" so they want to figure out or they want us to figure out this height right over here.

We need to figure out this height. And as we see, assuming that this is a normal building and it's built at a right angle to the ground, this triangle formed by the ladder the building in the ground is a right triangle. So the Pythagorean theorem will apply.

And the Pythagorean theorem tells us that the square or the sum of the squares of the two shorter sides is going to be equal to the square of the longer side or the square of the hypotenuse and the longest side is a side opposite the 90-degree angle and that's the hypotenuse.

So this tells us - the Pythagorean theorem tells us - that 8 squared plus h squared - H for height - is going to be equal to 60 squared. And 8 squared is 64. Sixty-four plus h squared is equal to 3,600. Subtract 64 from both sides - so let's subtract 64 from both sides - and we get h squared - h squared is going to be equal to - what is this - 3536 - and then this doesn't pop out into my brain is some type of perfect square so let's take a calculator out.

So let's try it out. So we want to take the square root of 3536 - second square root - you see a little square root symbol there in orange just want to press the orange button first. Square root of three thousand five hundred and thirty-six.

And I get fifty-nine point four six. And they want us to round our answer to the nearest tenth. So 59 point four six is greater than or equal to five so round up - so it rounds up to 59.5.

So H is going to be equal to maybe I should say approximately equal to - I already forgot the number - fifty-nine point five. Fifty-nine point five. And we're done!

Appendix B. Competence IAT: Description and Development

Implicit stereotypes. To measure implicit racial stereotypes, I developed an implicit association test relating race (Black/White) and competence. In its original form, the traditional Black/White IAT is a valence measure, meaning that it measures the relative strength of one's pairing of a positive versus negative valence with White people versus with Black people. It does this through a computerized timed classification task that compares how quickly and accurately test-takers can classify stimuli representing White people (e.g., photographs of faces) when the race category is paired with a good vs. bad valence term (e.g., "joy" vs. "hurt") to how quickly and accurately they can classify stimuli representing Black people.

For my competence IAT, I use the categories "African American" and "European American," following the traditional IAT. Following Fiske et al. (2002), I use the categories "Competent" and "Incompetent," and the competence target words "intelligent," "confident," "capable," and "efficient." I use the incompetence target words "disorganized," "unqualified," "stupid," and "unskilled" (inspired by Vitriol, Ksiazkiewicz, & Farhart [2018]). The stimuli included photographs of Black and White adolescents (4 male, 4 female for each racial group), obtained through Getty images and piloted on Amazon's MTurk platform to ensure that the age and race of the photographed subjects were perceived as intended. Using a selection of photographs in which the perceived race was as intended, and in which subjects' perceived ages were similar across races/genders, I built the competence IAT using the iatgen online software (Carpenter et al., 2018).

Prior to Study 1, I tested whether the competence IAT differed from the traditional Black-White valence IAT by conducting a pilot on MTurk in which respondents (target sample of n=300; 40 dropped for excessive speed, yielding final n=260) were randomly assigned to

complete my competence IAT or the traditional Black/White IAT (respondents were paid \$1.00). In the pilot, internal consistency (based on split-half with Spearman-Brown correction) was .86 for the competence IAT and .85 for the traditional IAT (error rates were .086 and .091 for competence and traditional, respectively). A t-test showed that scores on the traditional and competence IAT were significantly different ($p=.013$). On both tests, the average respondent showed significant pro-White bias, though the magnitude was smaller for the competence IAT (average traditional d -score = .42; average competence d -score = .30).

Validity Evidence for Black/White competence IAT. As validity evidence for my implicit measure, I fit a series of regression models predicting individuals' competence IAT d -scores (divided by the sample SD). To establish known-groups validity, I fit a series of models with demographic variables (indicators for race/ethnicity, gender, and whether the respondent worked in the field of education). To establish convergent validity, I fit models with academic expectations (as measured by the graduation guess item), gap prioritization, and gap explanations (or PCA indices) as the predictors. Results are presented in Table B1.

In Table B2, I present the correlations of the competence IAT scores with other relevant survey items from the Study 1 sample. Demonstrating initial validity evidence, respondents' competence IAT d -scores were significantly correlated with several survey items. Correlations were small but similar to the average of .12 found in a meta-analysis of implicit and explicit racial attitudes (Greenwald, Poehlman, Uhlmann, & Banaji, 2009). Respondents with more pro-White implicit competence bias showed lower guesses for Black students' high school graduation rates ($r = -.094$), gave less priority to closing racial achievement gaps ($r = -.185$), were less likely to believe that school quality played a larger role in racial achievement inequality ($r = -.175$), and were less likely to believe discrimination and racism played an

important role in racial achievement inequality ($r = -.154$). People showing more pro-White competence bias on the IAT may also be more likely to believe that parenting plays an important role in racial achievement disparities ($r = .085, p < .10$) and may be less likely to believe that income plays an important role ($r = -.08, p < .10$). Implicit bias did not predict the extent to which respondents believed that motivation, genetics, neighborhood, or home environment helped explain racial achievement disparities.

APPENDICES: EFFECTS OF ACHIEVEMENT GAP NEWS REPORTING

Table B1.
 OLS Regression Models Predicting Implicit Competence Stereotypes (Controlling for Treatment Condition).

	(1)	(2)	(3)	(4)	(5)	(6)
	D-score (std)	D-score (std)	D-score (std)	D-score (std)	D-score (std)	D-score (std)
Educator	-0.292* (0.144)	-0.217 (0.134)			-0.208 (0.133)	-0.201 (0.132)
AG	0.200* (0.0876)	0.224** (0.0818)	0.200* (0.0877)	0.170~ (0.0868)	0.201* (0.0820)	0.178* (0.0814)
Black		-1.034*** (0.113)			-0.972*** (0.113)	-0.982*** (0.113)
Latinx		-0.313~ (0.187)			-0.379* (0.184)	-0.369* (0.183)
Asian		-0.185 (0.175)			-0.211 (0.173)	-0.200 (0.172)
Other Race		-0.831 (0.653)			-0.791 (0.645)	-0.827 (0.641)
Multi-racial		-0.664*** (0.157)			-0.661*** (0.156)	-0.640*** (0.155)
American Indian		-0.984* (0.461)			-0.764~ (0.458)	-0.849~ (0.452)
Female		0.0186 (0.0899)			0.0877 (0.0900)	0.0704 (0.0886)
Non-binary		-0.894* (0.420)			-0.696~ (0.420)	-0.748~ (0.414)
Grad Guess			-0.00329 (0.00202)	-0.00356~ (0.00201)	-0.00430* (0.00189)	-0.00463* (0.00188)
Gap Priority (std)			-0.105* (0.0514)	-0.120* (0.0502)	-0.113* (0.0477)	-0.123** (0.0467)
Sch. Quality			-0.139* (0.0553)		-0.113* (0.0519)	
Motivation			-0.0232		-0.0118	

APPENDICES: EFFECTS OF ACHIEVEMENT GAP NEWS REPORTING

Parenting			(0.0491)		(0.0455)	
			0.115*		0.114*	
			(0.0569)		(0.0528)	
Discrimination			-0.0754		-0.0237	
			(0.0477)		(0.0452)	
Genetics			0.0444		0.0275	
			(0.0333)		(0.0315)	
Neighborhood environment			-0.0203		-0.0245	
			(0.0581)		(0.0542)	
Home environment			0.0219		-0.000807	
			(0.0681)		(0.0637)	
Income			0.0358		0.0497	
			(0.0511)		(0.0479)	
Explanation index (std)				-0.0239		0.0199
				(0.0491)		(0.0459)
Non-structural index (std)				0.164***		0.103*
				(0.0452)		(0.0427)
Constant	0.386***	0.613***	0.738*	0.564***	0.768**	0.838***
	(0.0637)	(0.0931)	(0.291)	(0.129)	(0.278)	(0.141)
<i>N</i>	514	514	514	514	514	514
<i>R</i> ²	0.018	0.181	0.088	0.074	0.234	0.223
<i>F</i>	4.681	11.09	4.396	8.139	7.541	10.21

Standard errors in parentheses. Outcome is d-score divided by its SD. AG=randomly assigned to the achievement gap news video (versus counter-stereotypical video). Demographic variables are binary indicators for the named category (White is the omitted racial group, male is omitted gender). Grad Guess = guess of Black HS graduation rate; Gap priority = how much of a priority believes is to close academic achievement gap between Black and White students (standardized from original 1=not a priority to 5=essential). Gap explanations items give extent to which respondent believes that factor is responsible for racial academic achievement gap between Black and White students (1=not at all; 2=slightly; 3=somewhat; 4=quite; 5=extremely). Explanation index = PCA index positively weighting all explanations for achievement gaps; Non-structural explanation index = PCA index positively weighting non-structural explanations for gap.

~ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table B2.
Correlations of IAT competence d-scores with other outcome variables (Study 1)

	IAT d-score
Graduation rate guess	-0.09 [*]
Gap Priority	-0.19 ^{***}
Explanation index	-0.06
Non-structural explanation index	0.21 ^{***}
Explanation: school quality	-0.18 ^{***}
Explanation: student motivation	-0.01
Explanation: parenting	0.08 [~]
Explanation: discrimination & racism	-0.15 ^{***}
Explanation: genetics	0.07
Explanation: neighborhood environment	-0.05
Explanation: home environment	0.01
Explanation: income	-0.08 [~]

[~] $p < 0.10$, ^{*} $p < 0.05$, ^{**} $p < 0.01$, ^{***} $p < 0.001$

Note. Graduation rate guess = guess of Black HS graduation rate; Gap priority = how much of a priority believes is to close academic achievement gap between Black and White students (1=not a priority to 5=essential). Explanation index = PCA index positively weighting all explanations for achievement gaps; Non-structural explanation index = PCA index positively weighting non-structural explanations for gap; Gap explanations items give extent to which respondent believes that factor is responsible for racial academic achievement gap between Black and White students (1=not at all; 2=slightly; 3=somewhat; 4=quite; 5=extremely)

Appendix C. Achievement Inequality Prioritization and Explanations.

In Study 1, I used several items taken or adapted from Valant & Newark (2016), who in turn adapted items validated by Feldman & Huddy (2010) (whose items came from the General Social Survey). I measured the extent to which respondents prioritized racial achievement disparities with the item, “As you may know, there is a racial academic achievement gap between Black and White students in the US. Thinking about all of the important issues facing the country today, how much of a priority do you think it is to close the racial academic achievement gap between Black and White students?” Answer choices were on a 5-point scale (1=not a priority; 2=low priority; 3=medium priority; 4=high priority; 5=essential).

I then surveyed respondents on their beliefs about the sources of racial achievement disparities with the item, “To what extent do you believe each of these factors is responsible for the racial academic achievement gap between Black and White students?” Respondents were asked to rate the contributions of the following possible explanations (with order randomized): School quality, student motivation, parenting, discrimination and racism, genetics, neighborhood environments, home environments, and income levels. Answer choices were on a 5-point scale (1= not at all; 2= slightly; 3=somewhat; 4=quite; 5=extremely; items were inspired by Valant & Newark [2016]).

I created two indices from the explanation items using principal components analysis (PCA). The PCA revealed two components with eigenvalues above 1. The first component (eigenvalue = 3.69) positively weighted all items and explained 46% of the total variation. The second component – which I call “non-structural” (eigenvalue = 1.21, explaining 15% of total variation) – positively weighted the motivation, parenting, genetics, and home environment explanations and negatively weighted the school quality, discrimination, and income

explanations (with a negative, but near-zero, weight for neighborhoods). As such, people who scored highly on this index tended to discount structural explanations for racial achievement disparities, in favor of cultural and genetic explanations. In Table C1, I present the PCA weights for each component. Videos did not have a significant effect on either index. In Table C2, I present the results for each item individually.

Table C1.

Weights from PCA for first and second principal components (Study 1; N=565)

Item	Explanations (component 1)	Non-structural explanations (component 2)
Sch. quality	.356	-.328
Motivation	.357	.298
Parenting	.360	.464
Discrimination & racism	.330	-.513
Genetics	.225	.292
Neighborhood enviro.	.401	-.048
Home enviro.	.407	.273
Income	.361	-.403

Note. See narrative text above for item wording

APPENDICES: EFFECTS OF ACHIEVEMENT GAP NEWS REPORTING

Table C2.
 OLS Regression Models Estimating Treatment Effects on Explanations for Achievement Gaps.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	School Quality	Motivation	Parenting	Discrimination	Genetics	Neighborhood	Home	Income
AG	0.142~ (0.0858)	0.135 (0.0934)	0.140 (0.0918)	0.0771 (0.0999)	0.00718 (0.120)	0.0210 (0.0857)	0.126 (0.0827)	-0.0805 (0.0935)
Black	0.460*** (0.119)	0.139 (0.130)	0.0917 (0.128)	0.619*** (0.139)	-0.0196 (0.167)	0.257* (0.119)	0.137 (0.115)	0.393** (0.130)
Latinx	0.0525 (0.195)	0.0977 (0.212)	0.188 (0.208)	0.203 (0.227)	0.586* (0.272)	0.0411 (0.194)	0.0228 (0.188)	0.0285 (0.212)
Asian	0.118 (0.181)	-0.118 (0.197)	0.0753 (0.194)	0.274 (0.211)	0.235 (0.253)	-0.0436 (0.181)	0.0410 (0.175)	0.117 (0.197)
Other Race	0.978~ (0.507)	0.592 (0.552)	0.641 (0.543)	0.220 (0.591)	0.910 (0.709)	-0.182 (0.507)	1.085* (0.489)	0.281 (0.553)
Multi-racial	-0.217 (0.171)	-0.305 (0.186)	-0.363* (0.183)	0.173 (0.199)	-0.519* (0.238)	-0.227 (0.170)	-0.362* (0.164)	-0.00271 (0.186)
American Indian	1.156* (0.507)	0.0249 (0.552)	0.321 (0.543)	1.182* (0.591)	-0.0931 (0.709)	0.308 (0.507)	-0.228 (0.489)	-0.429 (0.553)
Female	0.239** (0.0919)	-0.0250 (0.100)	0.113 (0.0983)	0.310** (0.107)	-0.397** (0.128)	0.174~ (0.0918)	0.203* (0.0886)	0.00522 (0.100)
Non-binary	0.484 (0.460)	-0.110 (0.501)	-0.0354 (0.493)	1.333* (0.536)	-1.740** (0.644)	0.785~ (0.460)	0.965* (0.444)	0.344 (0.502)
Constant	3.618*** (0.0942)	3.637*** (0.103)	3.768*** (0.101)	3.355*** (0.110)	3.036*** (0.132)	3.839*** (0.0942)	3.781*** (0.0909)	3.737*** (0.103)
<i>N</i>	565	565	565	565	565	565	565	565
<i>R</i> ²	0.060	0.016	0.020	0.061	0.052	0.023	0.040	0.020
<i>F</i>	3.913	0.997	1.286	3.980	3.368	1.473	2.545	1.260

Standard errors in parentheses. Item: “As you may know, there is a racial academic achievement gap between Black and White students in the US. Thinking about all of the important issues facing the country today, how much of a priority do you think it is to close the racial academic achievement gap between Black

APPENDICES: EFFECTS OF ACHIEVEMENT GAP NEWS REPORTING

and White students?” Answer choices were on a 5-point scale (1=not a priority; 2=low priority; 3=medium priority; 4=high priority; 5=essential). AG=binary indicator that respondent was randomly assigned to the Achievement Gap news clip condition (versus the counter-stereotypical condition). Demographic variables are binary indicators for the named category (White is the omitted racial group, male is omitted gender).

~ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix D. Subgroup Analyses (Study 1).

In Table D1, I present AGD effects from Study 1 on IAT *d*-scores by race/ethnicity, gender, and educator status. These analyses should be considered exploratory rather than confirmatory (and note small subgroup sizes in many cases). White respondents, as the largest subgroup, are the only subgroup who show a significant effect of being randomly assigned to the achievement gap news clip video (vs. the counter-stereotypical video; $ES=.27, p<.01$) Asian respondents show near-zero effect (though note the small sample size), and Black respondents are the only group with a negatively-signed coefficient ($ES= -.11, n.s.$). Female respondents ($ES=.248, p<.05$), but not male respondents ($ES=.137, n.s.$), show significant treatment effects, while educators ($n=53$) showed a near-zero effect estimate.

In Table D2, I present AGD effects from Study 1 on the Black high school graduation rate guess item by race/ethnicity, gender, and educator status. Again, White respondents and female respondents show significant effects of being assigned to the AGD video ($b= -5.34$ and -7.9 , respectively). The effect estimate for Black respondents is quite similar in magnitude to the effect estimate for White respondents ($b=-5.4$), though it is estimated much less precisely. Descriptively, the effect is smaller for male respondents ($b=-2.9$) than female respondents, and somewhat smaller for educators ($b=-2.97$) than non-educators. The large estimate among Asians ($b= -17.93$) should be interpreted cautiously given the small sample size.

APPENDICES: EFFECTS OF ACHIEVEMENT GAP NEWS REPORTING

Table D1.

OLS Regression Models Estimating Effects on IAT Competence d-score by Subgroup (Study 1)

	(1) Black	(2) White	(3) Latinx	(4) Asian	(5) Multi-racial	(6) Female	(7) Male	(8) Educator
AG	-0.112 (0.191)	0.271** (0.101)	0.175 (0.351)	0.00774 (0.357)	0.472 (0.339)	0.248* (0.103)	0.137 (0.166)	0.000741 (0.277)
Constant	-0.253~ (0.142)	0.572*** (0.0722)	0.297 (0.258)	0.511* (0.216)	-0.180 (0.206)	0.352*** (0.0724)	0.364** (0.116)	0.191 (0.194)
<i>N</i>	82	332	26	30	38	361	148	53
<i>R</i> ²	0.004	0.021	0.010	0.000	0.051	0.016	0.005	0.000
<i>F</i>	0.342	7.164	0.249	0.000471	1.937	5.776	0.686	0.00000717

Standard errors in parentheses. AG=binary indicator that respondent was randomly assigned to the Achievement Gap news clip condition (versus the counter-stereotypical condition).

~ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table D2.

OLS Regression Models Estimating Effects on Black Graduation Rate Guess by Subgroup Study 1)

	(1) Black	(2) White	(3) Latinx	(4) Asian	(5) Multi-racial	(6) Female	(7) Male	(8) Educator
AG	-5.395 (4.934)	-5.241* (2.159)	-13.93 (9.753)	-17.93* (8.474)	-7.546 (7.644)	-7.902*** (2.090)	-2.895 (3.619)	-2.974 (6.329)
Constant	54.85*** (3.623)	56*** (1.560)	59.31*** (7.245)	56.50*** (5.437)	54.76*** (4.580)	57.64*** (1.476)	52.08*** (2.595)	48.81*** (4.440)
<i>N</i>	89	366	29	34	39	383	177	63
<i>R</i> ²	0.014	0.016	0.070	0.123	0.026	0.036	0.004	0.004
<i>F</i>	1.196	5.893	2.041	4.477	0.975	14.30	0.640	0.221

Standard errors in parentheses. AG=binary indicator that respondent was randomly assigned to the Achievement Gap news clip condition (versus the counter-stereotypical condition).

~ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Appendix E. Follow-up Analyses with Mturk Sample (Study 2).

In Table E1, I present results for the graduation rate guess item from the two-week follow-up survey of Study 2 MTurk respondents. As can be seen in Columns 1 (no controls) and 2 (including demographic controls), the effects that were observed on the first survey were not sustained at follow-up. Given the sample attrition, it is possible that this apparent fade-out is due to heterogeneous initial effects. However, initial effects were observed among the sample of follow-up respondents (AG coefficient, Columns 3 and 4). An analysis of the change in graduation rate guess from the initial survey to follow-up survey (columns 5 and 6) suggests that the fading of effects at follow-up was driven by a rise in the graduation rate guesses among respondents who had been randomly assigned to the achievement gap (AG) video. Specifically, while respondents in the counter-stereotypical (CS) group did not show any change in their graduation rate guesses from survey 1 to follow-up (.48 percentage points), descriptively, AG respondents showed a greater rise in their graduation rate guesses (approximately 2 percentage points greater, *n.s.*). This suggests that AGD viewers may be returning to baseline after having their guesses influenced downward by the news clip.

In Table E2, I present descriptive statistics for the demographic and outcome collected from Study 2 MTurk respondents during the follow-up.

APPENDICES: EFFECTS OF ACHIEVEMENT GAP NEWS REPORTING

Table E1.
Follow-up Effect Estimates on Graduation Rate Guess (Study 2, Follow-up Respondents Only).

	(1) Grad Guess Follow-up	(2) Grad Guess Follow-up	(3) Grad Guess (1st)	(4) Grad Guess (1st)	(5) Guess Change	(6) Guess Change
AG	-3.319 (2.758)	-3.472 (2.813)	-5.383* (2.571)	-4.598~ (2.593)	2.065 (3.059)	1.126 (3.104)
CS	-1.103 (2.901)	-1.131 (2.939)	-2.192 (2.704)	-1.341 (2.709)	1.089 (3.217)	0.210 (3.242)
Black		0.716 (6.016)		5.881 (5.545)		-5.165 (6.637)
Latinx		8.504 (6.771)		7.741 (6.242)		0.762 (7.470)
Asian		0.159 (4.355)		5.359 (4.015)		-5.200 (4.805)
Other Race		1.651 (4.511)		-6.725 (4.158)		8.376~ (4.976)
Female		2.540 (2.353)		1.790 (2.169)		0.750 (2.596)
Constant	63.87*** (2.072)	62.51*** (2.365)	63.39*** (1.932)	61.73*** (2.180)	0.479 (2.298)	0.785 (2.609)
<i>N</i>	237	237	237	237	237	237
<i>R</i> ²	0.007	0.018	0.019	0.052	0.002	0.024
<i>F</i>	0.774	0.602	2.261	1.788	0.228	0.797

Standard errors in parentheses. AG, CS=dummy variables for achievement gap narrative and counter-stereotypical videos, respectively (Pythagorean Theorem video group is the omitted reference). Sample sizes by condition: PT=71, CS=74, AG=92. Demographic variables are binary indicators for the named category. “Other Race” includes respondents who identified as American Indian, multi-racial, or “other race” (collapsed due to 0 or too few observations in one or more group).

~ $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

APPENDICES: EFFECTS OF ACHIEVEMENT GAP NEWS REPORTING

Table E2.
Descriptive Statistics for Study 2 Follow-up Respondents.

	Full Sample			Control			CS			AG			p
	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	
Asian	0.076		237	0.099		71	0.054		74	0.076		92	0.602
Black	0.038		237	0.042		71	0.027		74	0.043		92	0.839
Latinx	0.030		237	0.042		71	0.027		74	0.022		92	0.739
White	0.785		237	0.803		71	0.811		74	0.750		92	0.582
American													
Indian	0.008		237	0.000		71	0.014		74	0.011		92	0.642
Multi-racial	0.055		237	0.014		71	0.068		74	0.076		92	0.193
Other race	0.008		237	0.000		71	0.000		74	0.022		92	0.206
Female	0.401		237	0.366		71	0.392		74	0.435		92	0.666
Non-binary	0.000		237	0.000		71	0.000		74	0.000		92	N/A
Age 18 or lower	0.004		237	0.000		71	0.014		74	0.000		92	0.334
Age 19-29	0.295		237	0.310		71	0.257		74	0.315		92	0.681
Age 30-39	0.359		237	0.380		71	0.378		74	0.326		92	0.710
Age 40-49	0.165		237	0.141		71	0.189		74	0.163		92	0.737
Age 50-59	0.089		237	0.070		71	0.081		74	0.109		92	0.672
Age 60-69	0.072		237	0.085		71	0.068		74	0.065		92	0.883
Age 70-79	0.017		237	0.014		71	0.014		74	0.022		92	0.899
Graduation Rate Guess (original)	60.620	16.362	237	63.394	16.377	71	61.203	17.655	74	58.011	14.993	92	0.107
Graduation Rate Guess (follow-up)	62.241	17.443	237	63.873	16.594	71	62.770	18.193	74	60.554	17.509	92	0.463
Change (original to follow-up)	1.620	19.300	237	0.479	19.349	71	1.568	19.995	74	2.543	18.855	92	0.796

Note. P-value is for F-test of equality across conditions. Rows without SD are binary indicators for named row category.

APPENDICES: EFFECTS OF ACHIEVEMENT GAP NEWS REPORTING

Appendix F. Additional Detail on Study 3.

In Study 3, I use items adapted from the General Social Survey (GSS; Smith, Marsden, & Hout, 2015) to measure racial stereotypes. While the question item stems mirror those used in the GSS, I adapted the scale to reflect recommended best practices in survey development (Gehlbach & Brinkworth, 2011). First, the original GSS items use a 1-7 scale but only provide substantive descriptors to the values of 1, 4, and 7. I added substantive descriptors to numbers 2, 3, 5, and 6 (shown below). Second, the GSS included stereotype items for the hard-working/lazy continuum and the unintelligent/intelligent continuum. I added items for the incompetent/competent and incapable/capable continua in an effort to improve the reliability of the stereotype index. The text of all survey items is given below (the order of items within each set was determined randomly for survey-takers).

In Table F1, I present descriptive statistics and randomization balance for the Study 3 sample.

APPENDICES: EFFECTS OF ACHIEVEMENT GAP NEWS REPORTING

STEREOTYPE ITEMS

In what follows, you will see a scale on which the characteristics of people from different groups can be rated.

Do people in this group tend to be **hard-working** or tend to be **lazy**?

White Americans

Almost all are lazy (1)	Many are lazy (2)	Slight majority are lazy (3)	No tendency to one or another (4)	Slight majority are hardworking (5)	Many are hardworking (6)	Almost all are hardworking (7)
-------------------------	-------------------	------------------------------	-----------------------------------	-------------------------------------	--------------------------	--------------------------------

Do people in this group tend to be **hard-working** or tend to be **lazy**?

Black Americans

Almost all are lazy (1)	Many are lazy (2)	Slight majority are lazy (3)	No tendency to one or another (4)	Slight majority are hardworking (5)	Many are hardworking (6)	Almost all are hardworking (7)
-------------------------	-------------------	------------------------------	-----------------------------------	-------------------------------------	--------------------------	--------------------------------

Do people in this group tend to be **unintelligent** or tend to be **intelligent**?

White Americans

Almost all are unintelligent (1)	Many are unintelligent (2)	Slight majority are unintelligent (3)	No tendency to one or another (4)	Slight majority are intelligent (5)	Many are intelligent (6)	Almost all are intelligent (7)
----------------------------------	----------------------------	---------------------------------------	-----------------------------------	-------------------------------------	--------------------------	--------------------------------

Do people in this group tend to be **unintelligent** or tend to be **intelligent**?

Black Americans

Almost all are unintelligent (1)	Many are unintelligent (2)	Slight majority are unintelligent (3)	No tendency to one or another (4)	Slight majority are intelligent (5)	Many are intelligent (6)	Almost all are intelligent (7)
----------------------------------	----------------------------	---------------------------------------	-----------------------------------	-------------------------------------	--------------------------	--------------------------------

Do people in this group tend to be **incompetent** or tend to be **competent**?

White Americans

Almost all are incompetent (1)	Many are incompetent (2)	Slight majority are incompetent (3)	No tendency to one or another (4)	Slight majority are competent (5)	Many are competent (6)	Almost all are competent (7)
--------------------------------	--------------------------	-------------------------------------	-----------------------------------	-----------------------------------	------------------------	------------------------------

APPENDICES: EFFECTS OF ACHIEVEMENT GAP NEWS REPORTING

Do people in this group tend to be **incompetent** or tend to be **competent**?

Black Americans

Almost all are incompetent (1)	Many are incompetent (2)	Slight majority are incompetent (3)	No tendency to one or another (4)	Slight majority are competent (5)	Many are competent (6)	Almost all are competent (7)
--------------------------------	--------------------------	-------------------------------------	-----------------------------------	-----------------------------------	------------------------	------------------------------

Do people in this group tend to be **incapable** or tend to be **capable**?

White Americans

Almost all are incapable (1)	Many are incapable (2)	Slight majority are incapable (3)	No tendency to one or another (4)	Slight majority are capable (5)	Many are capable (6)	Almost all are capable (7)
------------------------------	------------------------	-----------------------------------	-----------------------------------	---------------------------------	----------------------	----------------------------

Do people in this group tend to be **incapable** or tend to be **capable**?

Black Americans

Almost all are incapable (1)	Many are incapable (2)	Slight majority are incapable (3)	No tendency to one or another (4)	Slight majority are capable (5)	Many are capable (6)	Almost all are capable (7)
------------------------------	------------------------	-----------------------------------	-----------------------------------	---------------------------------	----------------------	----------------------------

PRIORITY ITEMS

As you may know, there is a racial achievement gap between Black and White students in the US.

Thinking about all of the important issues facing the country today, how much of a priority do you think it is to close the racial achievement gap between Black and White students?

<input type="radio"/> Not a priority (1)	<input type="radio"/> Low priority (2)	<input type="radio"/> Medium priority (3)	<input type="radio"/> High priority (4)	<input type="radio"/> Essential (5)
--	--	---	---	-------------------------------------

How important is closing the Black/White achievement gap as a social justice issue?

<input type="radio"/> Not important (1)	<input type="radio"/> A little important (2)	<input type="radio"/> Somewhat important (3)	<input type="radio"/> Quite important (4)	<input type="radio"/> Extremely important (5)
---	--	--	---	---

How important is closing the Black/White achievement gap to the future of the United States?

<input type="radio"/> Not important (1)	<input type="radio"/> A little important (2)	<input type="radio"/> Somewhat important (3)	<input type="radio"/> Quite important (4)	<input type="radio"/> Extremely important (5)
---	--	--	---	---

APPENDICES: EFFECTS OF ACHIEVEMENT GAP NEWS REPORTING

How important is it that our national political leaders are committed to closing the Black/White achievement gap?

<input type="radio"/> Not important (1)	<input type="radio"/> A little important (2)	<input type="radio"/> Somewhat important (3)	<input type="radio"/> Quite important (4)	<input type="radio"/> Extremely important (5)
---	--	--	---	---

How urgent is it that we close the Black/White achievement gap?

<input type="radio"/> Not urgent (1)	<input type="radio"/> A little urgent (2)	<input type="radio"/> Somewhat urgent (3)	<input type="radio"/> Quite urgent (4)	<input type="radio"/> Extremely urgent (5)
--------------------------------------	---	---	--	--

APPENDICES: EFFECTS OF ACHIEVEMENT GAP NEWS REPORTING

Table F1
Descriptive Statistics and Randomization Balance for Study 3 Sample.

	Kahn Academy		News		p
	Mean	SD	Mean	SD	
<i>Sample Demographics</i>					
Latinx/Hispanic	0.115		0.109		0.807
Asian	0.084		0.118		0.169
White	0.666		0.664		0.978
Black	0.105		0.092		0.604
Other Race	0.030		0.016		0.258
Male	0.571		0.549		0.595
Nonbinary	0.007		0.010		0.676
Some HS	0.010		0.000		0.079
HS/GED	0.111		0.092		0.433
Some college	0.216		0.214		0.943
Associate's	0.108		0.112		0.884
Bachelor's	0.422		0.421		0.975
Master's	0.118		0.135		0.541
Doctorate	0.014		0.023		0.386
Other Education Level	0.000		0.003		0.324
Age: 18-29	0.284		0.299		0.676
Age: 30-39	0.382		0.391		0.808
Age: 40-49	0.209		0.145		0.038
Age: 50-59	0.071		0.122		0.035
Age: 60-69	0.051		0.039		0.509
Age: 70-79	0.003		0.003		0.985
<i>Priority items</i>					
Priority	3.520	1.202	3.615	1.166	0.327
Social Justice	3.730	1.247	3.822	1.222	0.358
Future of US	3.578	1.244	3.780	1.166	0.041
Nat'l Political Leaders	3.574	1.260	3.730	1.227	0.125
Urgent	3.375	1.311	3.543	1.261	0.111
Priority Index	3.555	1.150	3.698	1.104	0.122
Priority Index (Std)	-0.057	1.015	0.069	0.975	0.122
<i>Black stereotypes</i>					
Hardworking	4.807	1.375	4.727	1.350	0.470
Intelligent	4.794	1.356	4.622	1.307	0.114
Competent	4.892	1.346	4.753	1.293	0.199
Capable	5.132	1.333	5.128	1.256	0.974
<i>White stereotypes</i>					
Hardworking	5.206	1.219	5.138	1.111	0.476
Intelligent	5.172	1.199	5.128	1.084	0.637

APPENDICES: EFFECTS OF ACHIEVEMENT GAP NEWS REPORTING

Competent	5.243	1.133	5.227	1.162	0.862
Capable	5.453	1.063	5.461	1.131	0.930
<i>Mean stereotypes</i>					
Black	4.906	1.210	4.808	1.140	0.304
Black (Std)	0.047	1.032	-0.037	0.972	0.304
White	5.269	0.992	5.238	0.950	0.704
White (Std)	0.020	1.026	-0.011	0.982	0.704
White-Black Diff.	0.362	1.092	0.431	0.992	0.421
White-Black Diff. (Std)	-0.034	1.046	0.032	0.951	0.421
<i>N</i>		296		304	

Note. Items without SD listed are binary indicator variables for row category. *p*-value if for test of null hypothesis of equal group means. Std=standardized to mean=0 and SD=1

APPENDICES: EFFECTS OF ACHIEVEMENT GAP NEWS REPORTING

Additional References (Appendices)

- Feldman, S., & Huddy, L. (2010). *The structure of white racial attitudes*. Paper presented at the annual meeting of the American Political Science Association, Washington, DC. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1643879
- Fiske, S.T., Cuddy, A.J.C., Glick, P., & Xu, J. (2002) A model of (often mixed) stereotype content: Competence and warmth respectively flow from perceived status and competition. *Journal of Personality and Social Psychology*, 82, 878-902.
- Gehlbach, H., & Brinkworth, M.E. (2011). Measure twice, cut down error: A process for enhancing the validity of survey scales. *Review of General Psychology*, 15, 380-387.
- 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.
- Smith, T.W.; Marsden, P.V., & Hout, M. General Social Surveys, 1972-2014. [machine-readable data file]. Principal Investigator, Tom W. Smith; Co-Principal Investigators, Peter V. Marsden and Michael Hout, NORC ed. Chicago: National Opinion Research Center, 2015. 1 data file (59,599 logical records) and 1 codebook (3,485 pp)
- Valant, J., & Newark, D. A. (2016). The politics of achievement gaps: US public opinion on race-based and wealth-based differences in test scores. *Educational Researcher*, 45(6), 331-346.
- Vitriol, J.A., Ksiazkiewicz, A., & Farhart, C.E. (2018). Implicit candidate traits in the 2016 U.S. presidential election: Replicating a dual-process model of candidate evaluations. *Electoral Studies*, 54, 261-268.