




Using self-affirmations to improve achievement in fourth-grade students

Jisun Sunny Kim¹ · Thomas J. Kehle¹ · Melissa A. Bray¹ · Sierra M. Trudel¹  · Brenna Fitzmaurice¹ · Adeline Bray¹ · Marisa Del Campo¹ · Erin DeMaio¹

Accepted: 3 January 2022 / Published online: 14 February 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Disparity exists between racially minoritized students and their White student counterparts in academic achievement. This discrepancy engenders the difference with which students will have opportunities in advanced courses; rates of high school graduation, college acceptance, and completion rates; and salary and quality of life. The academic disparity between the two groups has been found to have roots in stereotype threat, which causes anxiety where the individual's behavior may confirm the negative stereotypes of one's in-group. Reducing stereotype threat has been theorized to allow minoritized students and those in negatively stereotyped groups to enhance their academic performance by removing levels of anxiety hampering their performance. Following previous work, whereby the academic achievement gap between Black and White middle school students were reduced, this study examines the effectiveness of such an intervention on 4th grade, elementary students' reading achievement levels.

Keywords Self-affirmation · Academic achievement · Intervention

The United States is considered one of freedom and equality, still the fact remains that a racial gap exists in academic achievement and performance between minoritized students and White students. Despite attempts to provide equitable education and opportunities for all individuals (e.g., *Brown vs. Topeka Board of Education, 1954*), persistent educational achievement gaps exist today across the U.S. (Hanushek et al., 2019). Evidence is clearly seen in results of standardized achievement assessments (National Center for Education Statistics [NCES], 2009, 2011, 2019), where on average Black fourth grade students performed 25 points lower (mean standard score of 224) than White peers on reading (mean scale score of 249) and 26 points lower on mathematics (mean standard score of 203) than White peers (mean standard score of 229 (NCES, 2019). The achievement discrepancy surfaces before children even enter into elementary school (Bowman et al., 2018; Fryer & Levitt, 2004), and widens over time (Bacharach et al., 2003; Yeung & Conley, 2008), with later consequences revealed

in disparities in graduation and dropout rates (DePaoli et al., 2018; Noguera & Akom, 2000).

Closing the achievement gap in the United States is crucial, as educational achievement is indisputably significant for increasing graduation rates, providing appropriate educational standards and expectations through establishing gifted and advanced placements among minoritized students (Milner, 2012; Olszewski-Kubilius et al., 2018), and most notably, for obtaining future economic growth (Appel & Kronberger, 2012; Hanushek et al., 2019). Stated simply, with the assistance of knowledge that stems from research, educators within the field of psychological and educational domain must utilize what is known about the external conditions of the educational learning process (i.e., school and classroom environment) to monitor and support (Kremenitzer, 2005) a child's development. This may be accomplished by first providing our students with equitable opportunities to produce work reflective of their capabilities, as well as encouraging them to reach beyond what others falsely perpetuate is their reach, via interventions such as self-affirmations.

✉ Jisun Sunny Kim
drj@urlifeinspired.com

¹ Department of Educational Psychology, University of Connecticut, Storrs, CT, USA

Explanations for Achievement Gap

Underperformance is often misattributed to deficiencies in the students and families (Howe et al., 1998), perpetuating the illusion that academic performance is determined by innate characteristics. This belief has negative consequences for students by reducing access to support and encouragement. Moreover, it contributes to a lack of progress and positive reform within educational systems by removing the responsibility to make necessary systemic changes in our schools and society (Comeaux & Jayakumar, 2007).

In contrast, research has furthered the understanding of how achievement gaps present and widen as a result of differences in access to opportunities (e.g., Mark, 2013; Milner, 2012). A student's access to opportunities depends on factors such as socioeconomic status, school quality, community resources, parental involvement, exposure to educationally enriching experiences, and teachers' expectations that foster or inhibit educational outcomes (Oates, 2009). There is great disparity in the types and number of opportunities that are accessible to students in the U.S., with minoritized students having disproportionately limited access. However, the systemic injustices that exist to perpetuate opportunity gaps have been difficult to remediate. Furthermore, while access to resources and opportunities plays a major role in setting students up for future success, these factors do not appear sufficient in explaining achievement differences when students are matched on these characteristics (Aronson, 2004; Steele & Aronson, 1995).

To comprehensively understand discrepancies in performance, an ecological perspective (Bronfenbrenner, 1977) is critical, as research shows that both the person's environment and characteristics interactively affect behavior (Stephens et al., 2012), and thus outcomes. That is, efforts to reduce achievement gaps must encompass comprehension of the individual's environment as well as the internal cognitive, emotional, and social processes that are ongoing for the individual functioning within that environment. While it is crucial to address systemic injustice and disparities, it is equally necessary to explore intrinsic variables, which often may be more efficiently modified through intervention.

Stereotype Threat: An Ecological Perspective

The study of stereotype threat has emerged as a compelling factor in explaining the achievement gap through an ecological approach (e.g., Appel & Kronberger, 2012;

Aronson, 2004; Bancroft et al., 2017; Hartley & Sutton, 2013; Kellow & Jones, 2008; Smith & Hung, 2008; Steele & Aronson, 1995; Taylor & Walton, 2011). Stereotype threat is defined as an inhibiting cognitive process that prevents individuals from performing to potential due to invasive thoughts about confirming a negative stereotype (Liu et al., 2021; Steele et al., 2002). To protect the ego, the individual attributes poor performance to factors that will not threaten sense of self-worth. In this sense, the concept shares similarities with that of self-handicapping (Spencer et al., 2016).

The Study of Stereotype Threat: Pathways and Characteristics

Stereotype threat affects performance through direct and indirect pathways (Owens & Massey, 2011). Specifically, it presents a direct pathway of influence by increasing cognitive load and thereby reducing potential. During this process, working memory is impaired as the individual responds to the stress of the threat, by hypermonitoring performance or suppressing uncomfortable emotions. Simultaneously, through an indirect pathway, individuals discredit the importance of academic performance in order to protect their self-esteem and sense of accomplishment. Motivation is reduced as a result of diminishing the significance of certain achievements as related to self-worth. It is thus hypothesized that individuals effectively remove the psychological stress of confirming the stereotype threat (Owens & Massey, 2011).

Studies have been conducted in experimentally controlled and contrived settings (e.g., Howard & Borgella, 2018), as well as in natural settings using surveys and latent variable structural equation models to unearth consistent findings (Owens & Massey, 2011). Many have utilized priming as part of the methodology to show the influence of stereotypes on individuals' thoughts and behaviors (Cheryan et al., 2009; Howard & Borgella, 2018). For example, older adults that were primed with elderly stereotypes before being shown a video of a crime had greater difficulty recalling details of the crime (Rossi-Arnaud et al., 2018); another study showed that implicit priming of negative gender-stereotypes through the use of a gaming leaderboard led women (ages 18–38; $m = 22.8$) to expect worse performance from themselves than women in the control condition (Vermeulen et al., 2016).

Racial Identity and Stereotype Threat Understanding the influence of stereotype threat on minoritized students' achievement requires familiarity with the development of racial-ethnic identity, or the extent to which an individual finds connections to a group based on shared social and physical traits and customs (Umaña-Taylor et al., 2014; Umaña-Taylor et al., 2014). One's group identification is largely dependent on both social and developmental

influences (Corenblum & Armstrong, 2012) beginning in childhood. Children are able to identify their own race by five to six years of age (Byrd, 2012); at the same time, they begin to show awareness of stereotypes and are able to understand that others may endorse different stereotypes from their own (Desert et al., 2009; Tomasetto et al., 2011). As children develop awareness of stereotypes, their vulnerability to their effects increases.

Cognitive maturity allows for stronger racial-ethnic identity (Umaña-Taylor et al., 2014), and the more an individual derives a sense of identity from a group, the more susceptible they become to a stereotype threat impacting that social group (Wout et al., 2008). The existence of subtle and overt biased attributions of negative characteristics to minoritized groups cannot be denied, a fact which has serious implications for achievement and other outcomes (Thoman et al., 2013; Walton et al., 2012). Racial priming, or bringing one's own race into awareness, occurs daily in the lives of stereotyped individuals through interactions with teachers, peers and colleagues, physical surroundings, media, and society (Owens & Massey, 2011; Umaña-Taylor et al., 2014). This can act as a reminder of existing stereotypes and can perpetuate false negative beliefs within individuals.

For minoritized populations, simply being underrepresented can induce a feeling of lack of belonging (Murphy et al., 2020; Walton & Cohen, 2007), which increases the chance of one perceiving a stereotype threat (Cook et al., 2012). Uncertainty of belonging and of the quality of social connectedness contributes to racial disparities in achievement (Murphy et al., 2020; Walton & Cohen, 2007); a meta-analysis of current literature on school belonging and student functioning found that higher school belonging relates to lower dropout rates in secondary students (Korpershoek et al., 2020). Students who do not feel like they belong experience low levels of motivation to follow rules in the schools and when prompted, 25% of students attributed not feeling like they belong to the reason why they dropped out of school (Juvonen, 2007).

Stereotype Threat in Academic Settings Stereotype threat effects are prevalent in academic settings and stereotyped students have heightened sensitivity to stereotypical cues in the classroom (Cook et al., 2012). They are especially salient in evaluative situations (Appel & Kronberger, 2012; Desert et al., 2009); simply indicating one's race, gender, or even student-athlete status activates stereotype threat before testing (e.g., Adams et al., 2006; Kellow & Jones, 2008; Riciputi & Erdal, 2017) and affects an individual's sense of competence, feelings of belonging, and trust in the people around him or her (Aronson, 2005; Mello et al., 2012). Some moderators of effects on performance include the meaning attached to the test (i.e., how one interprets or values the test and its results), beliefs about intelligence, and the salience of

social identity (Jordan & Lovett, 2007; Spencer et al., 2016). Stereotypes pertaining to performance as related to one's identity trait (e.g., gender, race) have been shown to have a self-fulfilling effect on achievement (Hartley & Sutton, 2013). For example, when women were primed with stereotypes about gender differences in mathematics performance, they exhibited lower performance and difficulty encoding critical math information. Similarly, white students' athletic performance will diminish when told athletic ability is an inherent trait (Hartley & Sutton, 2013). On the other hand, when a member of one's group is recognized for great endeavors that defy stereotypes, this provides a positive role model that can influence how individuals perceive their own behavior and abilities. For example, research has shown that Black individuals performed as well as White individuals on assessments when Obama's accomplishments were salient in their awareness (Marx et al., 2009).

Stereotype threat effects have also been studied in the context of effects on the process of learning itself (i.e., encoding information into memory to apply in future tasks; Rydell et al., 2010; Spencer et al., 2016). It is shown to impair both verbal working memory and sequential reasoning in the classroom (Appel et al., 2011). In one study, Black students who studied under stereotype salient conditions performed worse than students who studied in non-threatening conditions (Taylor & Walton, 2011).

While research supports the negative influence of stereotype threat on the performance of older students, less is presently known about effects on younger students. Research suggests that early events (e.g., transitions from elementary school to middle school) leave a lasting impression on students; when damaging notions are reinforced over time, later efforts to intervene have limited success in undoing damage (Cook et al., 2012). Thus, it is important to address the impacts of cognitive development and social understanding as factors contributing to stereotype's effects. Clarifying these processes would be beneficial for timing interventions to occur during potentially critical periods in development, which could increase the probability of their success (Ambady et al., 2001; Liu et al., 2021).

Self-Affirmation Interventions for Alleviating Stereotype Threat

To address educational underperformance, researchers have turned to psychosocial and mind-body health interventions, which attend to social and emotional aspects of an individual such as sense of belonging and mindset. Self-affirmation is a mind-body health and psychosocial intervention that relies on the premise that protecting and enhancing one's self-integrity and sense of worth

is a basic motivation (Klein et al., 2011; Martens et al., 2006); in addition, individuals can overcome a threat in one domain by affirming self-worth in other domains (Schmader & Johns, 2003), which allows for preservation of self-integrity. Self-affirmation interventions prompt the individual to consider and affirm aspects of self that enhance self-integrity. When faced with a stereotype threat, the self-affirmation intervention restores self-integrity by extending the domains of self-concept beyond the threatened domain (Critcher & Dunning, 2015). The act of self-affirmation is shown to lower levels of self-protecting behaviors, reduce defensiveness in processing self-relevant information, reduce physiological stress responses, and improve academic performance (Klein et al., 2011; Schmader & Johns, 2003). Mental health benefits include higher levels of self-control, and feelings of love, compassion, and connectedness (Nelson et al., 2014).

A simple self-affirmation intervention administered in middle school was shown to increase the grade point average (GPA) of Black students (Cohen et al., 2006; Cohen et al., 2009) and Hispanic students (Sherman et al., 2009) who were tracked for one to three years, effectively closing the achievement gap in their school populations. Implementation of a self-affirmation exercise has been shown to improve the math performance of women under stereotyped conditions, whereas performance of men did not change (Martens et al., 2006). Self-affirmation has been shown to influence outcomes for individuals with low self-esteem; affirmed individuals were not observed to psychologically distance themselves from their partners when faced with threat (as compared to individuals who did not receive self-affirmation interventions).

Addressing the achievement gap through the use of self-affirming, psychosocial intervention has repeatedly been demonstrated as an effective means for halting the expansion of the achievement gap that exists between minoritized students and their same-aged white peers. Social connectedness and group identity are core components of one's understanding and acceptance of self, and the use of self-affirmations in the context of education has been studied to address the achievement gap with respect to minoritized students, especially over the last two decades. Stereotypes that exist amongst American culture pose a psychological threat to Black and Hispanic communities overall, but can have detrimental effects to the academic growth and success of minoritized students. Findings from past and current research have consistently demonstrated the power of self-affirmation, with an emphasis on its applicability to minoritized populations in K-12 education. (Binning et al., 2021; Cohen & Sherman, 2014; Walton et al., 2012; Yeager et al., 2013).

Purpose of the Study

Cohen et al. (2006, 2009) conducted research to see the effects of psychosocial interventions (i.e., self-affirmation intervention) on performance of minoritized students by removing the stereotype threat. Their seminal work in the field of educational and social psychology utilizing self-affirmations have produced results indicating the significant positive effect of the administration of the intervention on middle-school students' academic performance. Participants consisted of middle- to lower-socioeconomic status families. Cohen et al. (2006, 2009) implemented randomized field experiments and a longitudinal follow-up study to measure the effectiveness of self-affirmations on academic achievement. This study aimed to extend current literature on the effects and values of self-affirmation as an intervention in the education system. Specifically, it looked to see if a study with students in primary grades (i.e., 4th grade), as opposed to older student populations, would reproduce such inspiring results as prior seminal studies by Cohen et al. (2006, 2009). The development of social and emotional skills in the early stages of a child's life is critical in setting up the foundation upon which future successes of the child is built (Denham et al., 2012; Kremenitzer, 2005; Pekrun, 2006) – which makes the purpose of this study, critical. Similar results would further corroborate what is believed to be the effects of stereotype threat in academic underperformance in minoritized students, as well as indicate how early the pervasive influence of stereotype threat begins.

The following research questions were proposed and directed this study: (1) will self-affirmation interventions designed to improve the academic achievement of minoritized students prove effective among 4th grade students, as it has been shown in previous studies among older, adolescent students (i.e., 7th and 8th grade, as well as college students)?; (2) are there differential effects of this intervention by subgroups of students (i.e., Black and Hispanic students)?; (3) will students in the self-affirmation condition illustrate a significant difference in their emotions related to academic settings (i.e., boredom, anxiety, enjoyment) as compared to students of the control condition?

Initially, this study sought to replicate previous studies and their results. Specifically, following the seminal work of Cohen et al. (2006, 2009), the study was proposed to assess for differential effects of treatment between minoritized students and White students (i.e., reduce the academic achievement gap between minoritized students and White students). However, as the study's sample did not include enough of a White population to answer this question statistically, the first research question was altered to reflect this prior to the administration of the intervention.

Instead, this study sought to explore differential effects of the intervention between Hispanic and Black students on their reading achievement scores.

Method

Participants

The sample included 69 fourth grade students attending a public elementary school in a southern Connecticut city, which was measured as performing in the lowest 10th percentile of schools on the state achievement test. The majority of students were Hispanic and Black in race, with only a couple students whom the school identified as White ($n=2$). All students were included in this study (i.e., there were no exclusionary criteria used in recruiting students); of the 88 fourth-grade students in the three classes, only 8 students' parents declined to participate by indicating on the permission forms that they did not wish for their child to participate in the study, and 11 student's parents failed to return the permission slips to the school after multiple attempts by teachers to return them; 3 students were removed from the study as they relocated to a different district and school during the time of study. As indicated by their teachers, all participants were English-speaking students (see Table 1 for demographic information).

A power analysis was conducted to determine the minimum number of participants required to provide statistically reliable results (i.e., increase probability of finding a significant difference when it exists). The researchers set the power at 0.8, the alpha value to be 0.05, and the effect size to be 0.1 using the G*Power software, a free online power analysis program to help determine the size of sample needed. Results indicated that a minimum sample size of 114 was required in finding statistically significant and meaningful results. However, due to difficulty in obtaining interest from schools, the study progressed with the participants from the

school with a total of 69 out of 88 fourth-grade students (66 after attrition).

Additionally, only after the study progressed, it was established that primary research question one had to be adjusted. That is, the demographics of the fourth-grade students in the study consisted of Hispanic and Black students; only one student in the study was considered White. The research question, therefore, was adjusted to reflect this sample of participants. Specifically, the question, *will self-affirmation interventions designed to reduce the academic achievement gap between minoritized and White students prove effective among 4th grade students, as it has been shown in recent studies among older, adolescent students?* was adjusted to test for the effectiveness of self-affirmation intervention on Black and Hispanic students' achievement (per Cohen et al., 2006) in lieu of assessing for the gap in effectiveness of the intervention on achievement between *minoritized* students and White students. The adjusted and current research question therefore reads, *will self-affirmation interventions designed to reduce the academic achievement gap between older, adolescent minoritized and White students prove effective in improving achievement of minoritized, 4th grade students?*

Experimental Design

Design and Variables A randomized between-group experiment was conducted to determine the effects of a self-affirmation intervention on fourth-grade students' academic achievement, as measured by dependent variables of post-intervention standardized reading achievement scores, GPA, and students' affect (enjoyment, anxiety, and boredom) in relation to different academic reading settings/tasks (classroom, homework, test). In addition to the independent variable of experimental condition, predictor variables included students' pre-intervention reading achievement scores, race, gender, and timing of intervention (i.e. class).

Control and Treatment Conditions Students were randomly assigned to either the experimental/treatment condition or the control condition. The control condition required students to choose from a list of values, that which is least important to them, and participated in a short, written exercise describing why this value may be important to others – consistent with methods employed by previous studies (i.e., Cohen et al., 2006, 2009). The treatment condition included a self-affirmation intervention, whereby students chose from a list of the same values which was most important to them and participated in a short, written exercise describing why this value was important to them.

Internal Validity Exploratory analyses were conducted to assess for outcome differences by intervention date in order

Table 1 Demographic information

	Self-Affirmation		Control	
	<i>n</i>	%	<i>n</i>	%
Gender				
Female	20	54	16	55
Male	17	46	13	45
Race				
Hispanic	29	78	24	83
Black	8	22	4	14
White	0	0	1	3

to reduce the internal validity threat of history (i.e., changes in student attitudes or behavior arising from passage of time). In addition, the study analyses assessed for differential growth between the control and treatment conditions to address the threat of regression to the mean.

Measures

Benchmark Scores Reading scores from the Houghton Mifflin Harcourt (HMH) reading assessment were used as a marker of academic achievement. Benchmark scores were obtained from and averaged at two points pre-intervention and two points post-intervention. As Cohen et al. (2006, 2009) found that students with weaker initial academic performance saw greater gains from self-affirmation, the influence of prior achievement was examined as a potential moderator.

Report Card Grades Report card grades (i.e., grade-point averages) were included in the study to allow for an assessment of achievement differences among students as reflected through teachers' scores; this also enabled researchers to consider possible attenuating effects of effort that may be reflected in student grades.

Affect Related to Academics Students' responses on the Achievement Emotions Questionnaire-Elementary School (AEQ-ES) were explored to find differences in student affect as related to academic settings or tasks (i.e., class, homework, and taking exams). The AEQ-ES has evidence as a psychometrically valid and reliable measure ($.71 \leq \alpha \leq .93$) and it is cross-culturally relevant (Lichtenfeld et al., 2012). Specifically, it assessed for eight distinct emotions: classroom enjoyment, classroom anxiety, classroom boredom, homework enjoyment, homework anxiety, homework boredom, test enjoyment, and text anxiety. For this study, the AEQ-ES was revised (with permission granted from author) to assess students' opinions regarding reading, the subject relevant to this study, while the original assessed for math. Both the female and male versions were utilized.

Procedure

Intervention Packet Construction Intervention packets were constructed for both the treatment and control conditions. They were created to be nearly identical in appearance and consisted of three stapled pages. While information in the packets was replicated from Cohen et al.'s (2006, 2009) preliminary studies, wording was modified to be easily understood by a fourth-grade student.

On both sets of packets, the first page was a cover page with the student's name and identifying teacher and school. The second page included instructions that directed students

to look at a box with a list of values and circle the value that is most or least important to them (depending on assigned condition). Values included: athletic ability, being good at art, creativity, doing things on my own, music, focusing on what's happening now, being part of a community, my racial group, school club, family and friends, religion, being funny. Instructions to the treatment and control condition differed with respect to whether they should circle the most or least important, and whether they should focus on why the value was important to them, or why it might be important to someone else. Both conditions were instructed to focus on their thoughts and feelings, and to "not worry about spelling or how it is written." The third page included a reinforcement of the condition. That is, students were asked to indicate how much they agreed with the following four statements: "this value has influenced my life," "this value is an important part of who I am," "I care about this value," and "I try to live up to the value".

Implementation During the first visit to the elementary school, the researcher received a list of the students in all three of the classes. Each student was randomly assigned to a condition through the use of a random number generator in Microsoft Excel. Unlike the preceding studies (Cohen et al., 2006, 2009) teachers were not asked to implement the study themselves. Instead, the researcher went into all three classrooms with a script and set of procedures for implementing the intervention while the teachers stayed in the classroom at their desks.

Throughout administration of the intervention, many students asked if they could pick more than one value. As Cohen et al. studies (2006, 2009) included versions of choosing just one value as well as choosing multiple values, the researcher responded that they can choose more than one should they choose to do so. In addition, the majority of students struggled to understand the concept of a Likert-scale; that is, they did not understand without researcher assistance how to indicate their responses according to how much they agreed with the statement.

The researcher followed a checklist on which implementation fidelity was recorded, as well as following a script and order of implementation procedures that were to be followed. A script was read each time (i.e., the script was not read verbatim, rather, all important elements of the introduction/script were included) the intervention was implemented and before the packets were handed out. In this introduction, students were asked to raise their hands quietly if they had any questions and the researcher would approach them and answer them. Upon handing out the packets, the students were reassured that there were no right or wrong answers and to answer honestly. The time that students began was recorded and they were given 15 min to review the packet and follow instructions. After the duration of 15 min,

students were reaffirmed for their honest opinions and packets were collected. If students had any further questions after the packets were collected, the researcher answered them appropriately. A list of possible questions and responses were drafted prior to implementation and included the following: (1) Why are we doing this? Why are you here? – to which the researcher responded approximately, *I'm here to see if the writing exercise you are doing today helps students do differently in their classes*; (2) Do we have to do this? – to which the researcher responded approximately, *You do not have to do this if you do not want to. You have the right to change your mind – if you no longer wish to participate, please raise your hand after we begin*; (3) What if I answer incorrectly? – to which the researcher responded approximately, *There are no right or wrong answers. I just want to see your honest responses. Do your best in being honest and completing the exercise. If you have any questions about what to do, you can raise your hand and I'll come over to help*; (4) How long does my response have to be? – to which the researcher responded approximately, *You should aim to write a couple of sentences, as many as you need to answer the first question. There are a couple of questions at the end that you can just circle your answers for*; (5) Is this being graded? – to which the researcher responded approximately, *These questions are not being graded – in fact, your teacher will not even see your responses, so do your best and answer honestly*. The list of probable questions and responses were documented in the Fidelity Checklist completed by the researcher. After everything was collected and any questions answered, the researcher recorded the time that administration had concluded.

Students who were scheduled to make up the intervention were pulled from the classroom in a public alcove that included tables and chairs. The students were provided with the same script and procedures as when implemented with the whole classroom.

Upon the passing of 20 school weeks (i.e., half a school year), the AEQ-ES was implemented. Students were provided with the AEQ-ES packet and instructed to answer each question the best they can. They were again reassured that there were no right or wrong answers, and that their teachers would not see their responses. The researcher followed a script and procedures as written in the Fidelity Checklist. At this time, remaining student achievement scores (i.e., post-implementation HMH and GPA were also collected). The successful translation of [modified] evidence-based interventions to practice in the fourth-grade elementary setting enabled the researcher to implement the intervention with high fidelity as demonstrated through consistent, repeated usage of a fidelity checklist to ensure all variables were controlled in a structured manner throughout the implementation stage of the study. In addition to the fidelity checklist, a script and an orderly set of implementation procedures were routinely

utilized in an organized manner. Although unlikely due to the level of structured implementation by the researcher, there may have been a slight, subjective impact on student performance as a result of the researcher implementing the intervention as opposed to the classroom teacher, as demonstrated in the preceding studies (Cohen et al., 2006, 2009). Students generally find their classroom teacher, a familiar authority figure, to be trustworthy. Considering no former relationship/rapport was established between the independent researcher and the student participants, it could be argued that students may have felt wary of completing the assigned packet, and may have felt more open to the idea of completing the packet had it been distributed by a trusting adult, such as their classroom teacher.

However, the researcher accounted for fixed responses to theoretical questions likely to be posed by students upon both first glance and completion of the distributed packets, many of which stemmed from ongoing reassurance that the packets were not to be graded, their teachers would not see their responses, and their best effort is all that was necessary in order to perform well on the assignment. This level of reassurance by the researcher could likely absolve any feelings of wariness/mistrust that may or may not have been subjectively felt by the fourth grade students. Collectively, the degree of implementation fidelity is considered to be high, as reflected in the carefully constructed procedures the researcher utilized to implement the intervention in the fourth-grade classroom setting to account for the intervention itself, as well as hypothetical scenarios with controlled outcomes. No other program adaptations were made.

Statistical Analyses

Data were analyzed using multiple regression and MANCOVA analyses on SPSS. Two multiple regression analyses were conducted to assess for effects of the intervention on two indicators of achievement: literacy scores (i.e., GPA) and HMH scores (i.e., Post-HMH), with dummy variables created for categorical predictor variables of ethnicity and HMH scores prior to implementation. Analysis utilizing MANCOVA assessed for academic affect/emotions (i.e. AEQ-ES scores) as related to intervention condition (i.e., self-affirmation vs. control) and will include a covariate, GPA. Thus, outcomes of interest when comparing treatment and control conditions included student scores on HMH assessments, GPA, and AEQ-ES results. Per Cohen et al.'s studies (2006; 2009), any two-way interactions between racial group and experimental conditions were calculated via a full regression model with main effects computed for racial group, gender, and experimental condition.

With each regression analysis (two, one for HMH scores, one for GPA), assumptions were analyzed for violations of

independence of cases, homoscedasticity, and normal distribution of errors. To test for assumptions, the Durbin-Watson statistic was used to assess for independence of observations; the assumption of equal residuals for all predicted values of the dependent variable was assessed visually utilizing a homoscedasticity plot; normal distribution of errors was assessed visually with a histogram.

Results

Research Question 1

Will self-affirmation interventions designed to reduce the academic achievement gap between older, adolescent minoritized and White students prove effective in improving the achievement scores of minoritized, 4th grade students? Two multiple regression analyses were conducted to assess the effectiveness of self-affirmation interventions on improving achievement scores (i.e., literacy scores [GPA] and Post-HMH scores). Additional predictors included in the analyses were ethnicity and HMH scores prior to implementation (i.e., pre-HMH).

Zero Order Correlations Prior to conducting the regression analyses, a zero-order correlation was conducted and found *intervention* and *ethnicity* did not significantly covary with *GPA* or *Post-HMH*. However, *pre-implementation HMH scores* (i.e., HMH scores prior to intervention implementation) significantly correlated with both *GPA* ($r = .754$) and *Post-HMH* ($r = .769$). *Gender* was excluded from the analysis as it did not significantly correlate with either outcome variable and was not of primary relevance to research questions.

Assumptions Met In assessing the effects of intervention and ethnicity on reading achievement scores, regression analyses were conducted; prior to reviewing results, assumptions were tested and critical ones were met. The assumptions of independent errors, homogeneity of variance and linearity, and normality of residuals were all met. In addition, assessing for outliers revealed no outliers. To test for effects of the intervention on reading related emotions, a MANCOVA was run. Prior to reviewing the results, assumptions were tested. Data revealed that the assumption of homogeneity of variance was met, and no multivariate outliers were found. However, normal distribution of scores was violated and a few univariate outliers were found; for two reasons, the data were not transformed: 1. the MANCOVA is robust for violation of normal distribution, 2. violation of these assumptions leads to an increase in Type I error, but we did not find significant results.

Table 2 Model Summary and F Statistic, PostHMH

	R ²	F	Sig.
Model 1 (PreHMH)	.568	84.254	$p < .05$
Model 2 (PreHMH, Ethnicity)	.575	27.908	$p < .05$
Model 3 (PreHMH, Ethnicity, Intervention)	.576	20.691	$p < .05$

Results, Post-HMH The multiple regression model for predicting *Post-HMH* reading achievement scores (see Table 2) with *ethnicity*, *intervention*, and *Pre-HMH* as predictor variables significantly accounted for approximately 59% of the variance in *Post-HMH* scores ($F_{6,59} = 14.034$, $p < .001$, $R^2 = .588$; see Table 2). Subsequent interactions between *PreHMH* and *intervention*, and *PreHMH* and *ethnicity* were examined and revealed that the interaction between *Pre-HMH* and *intervention* was not significant, $\beta = .044$, $t(59) = .236$, $p > .05$; indicating that the influence of pre-intervention achievement scores on post-intervention scores did not depend on the intervention condition. Additionally, *intervention* was not found to be significant in predicting *Post-HMH*, $\beta = -.026$, $t(59) = -.168$, $p > .05$. That is, the intervention condition did not provide any significant value for predicting the *Post-HMH* scores above and beyond *Pre-HMH* scores.

Results, GPA The multiple regression analysis conducted to predict GPA utilizing the same three predictor variables (*ethnicity*, *intervention*, and *PreHMH*) was also significant, and predicted 61% of the variance in GPA ($F_{6,59} = 15.317$, $p < .001$, $R^2 = .609$). In examining interaction effects, the interaction between *PreHMH* and *intervention* was not significant in predicting *GPA*, $\beta = -2.647$, $t(59) = -1.622$, $p > .05$, indicating that the influence of *PreHMH* scores on *GPA* did not depend on *intervention* condition. With respect to the main effect of intervention, it was not found to be significant in predicting *GPA*, $\beta = 2.183$, $t(59) = 1.594$, $p > .05$.

A hierarchical regression analysis was conducted to see how much of a change (despite insignificance) the intervention provided in *GPA* scores, with *PreHMH* in the first model, *PreHMH* and *ethnicity* in the second model (as *ethnicity* had a higher correlation to *PostHMH*), and *PreHMH*, *ethnicity*, and *intervention* in the third model. The interaction effects were removed from hierarchical analysis because of non-significant results as indicated previously (the interaction between *PostHMH* and *ethnicity* was found to be not significant, as discussed below). The addition of *intervention* did not account for any change in variability in *GPA* (see Table 3).

Table 3 Model summary and F statistic, GPA

	R ²	F	Sig.
Model 1 (PreHMH)	.591	92.341	$p < .05$
Model 2 (PreHMH, Ethnicity)	.592	29.929	$p < .05$
Model 3 (PreHMH, Ethnicity, Intervention)	.592	22.089	$p < .05$

Research Question 2

Are there differential effects of this intervention by subgroups of students? Specifically, was there a significant difference between Hispanic students' and Black students' achievement scores as dependent on the treatment condition? To assess for the effects of ethnicity, two regression analyses were conducted using 0–1 coding to predict *PostHMH* and GPA. The analyses included *ethnicity*, *intervention*, and *PreHMH* consequent interaction variables (i.e., *PreHMH*ethnicity*, *PreHMH*intervention*). One participant's data was removed from the analyses as they were the only White student in the sample.

Results, PostHMH The interaction between *PreHMH* and *ethnicity* was not significant, $\beta = -.291$, $t(59) = -1.279$, $p > .05$, indicating that *PreHMH* did not variably affect *PostHMH* depending on the student's ethnicity. Additionally, there were no significant effects of *ethnicity* for predicting *PostHMH*, $\beta = .257$, $t(59) = 1.334$, $p > .05$. Upon further review, it was found that *ethnicity* only increased the percentage of variability in the dependent variable (i.e., predictive power added to the model by adding ethnicity to the model) by .002, or .2%.

Results, GPA The interaction between *PreHMH* and *ethnicity* was not significant, $\beta = -.307$, $t(59) = -.153$, $p > .05$. In addition, there were no significant effects of ethnicity for predicting GPA, $\beta = .331$, $t(59) = .195$, $p > .05$; *ethnicity* only increased the R² by .001, or .1% as compared to the first model (i.e., only *PreHMH* as a predictor).

Research Question 3

Will students in the self-affirmation condition illustrate a significant difference in affect/emotion as related to academic settings or tasks? The three specific emotions that were assessed by the AEQ-ES included enjoyment, anxiety, and boredom, as each occurs within three academic conditions: the classroom, with homework, and with tests. There were 23 subjects in the control condition and 30 subjects in the treatment condition (self-affirmation) that responded to the AEQ-ES to investigate this question. Thirteen students'

Table 4 Between subjects effects

	DV	F	Sig.	
GPA	Class Enjoyment	.270	.606	
	Class Anxiety	5.554*	.023	
	Class Boredom	1.242	.271	
	HW Enjoyment	.035	.852	
	HW Anxiety	3.299	.068	
	HW Boredom	.319	.575	
	Test Enjoyment	.882	.352	
	Test Anxiety	.744	.393	
	Intervention	Class Enjoyment	.456	.503
		Class Anxiety	.053	.818
Class Boredom		.706	.405	
HW Enjoyment		1.952	.169	
HW Anxiety		1.726	.195	
HW Boredom		1.625	.209	
Test Enjoyment		1.272	.265	
Test Anxiety		1.004	.321	

*Significant at the .05 level

responses were missing from the data because these children were not present during the time of the survey. A multivariate analysis of covariance (MANCOVA) was conducted to investigate the effects of the self-affirmation intervention in these domains, while controlling for GPA as a covariate. The interaction between GPA and *intervention* was not significant and therefore removed from analysis.

Testing for Assumptions A multivariate analysis of covariance was run to determine the effect of self-affirmation treatment on emotions related to reading. Preliminary assumption testing revealed that data were not normally distributed, as assessed by Shapiro-Wilks test ($p < .05$); however, as the MANCOVA is robust for this violation of assumption, analysis continued. Levene's test of equality of variance confirmed equal variance. There were some univariate outliers as assessed by inspection of boxplots – specifically, within test anxiety, class boredom, and class anxiety measures. However, no multivariate outliers were found as assessed by Mahalanobis distance ($p > .001$). There was homogeneity of variance-covariance matrices, as assessed by Box's M test ($p = .413$).

Results The Wilks' lambda was used to test the null hypotheses. Differences in reported emotions regarding academic situations (i.e., context of reading) were not statistically significant across GPA scores, $F_{8, 40} = 1.351$, $p > .05$; Wilks' $\Lambda = .787$; partial $\eta^2 = .213$; or between intervention conditions, $F_{8, 40} = .377$, $p > .05$; Wilks' $\Lambda = .930$; partial $\eta^2 = .070$. That is, students' reported emotional levels surrounding each of the eight emotion-situation categories

Table 5 MANCOVA descriptive statistics & independent samples test

	Intervention	Mean	SD	N	t	Sig.
Class Enjoyment	Control	3.595	1.150	21	-.880	.383
	Treatment	3.802	1.119	29		
	Total	3.715	1.125	50		
Class Anxiety	Control	1.548	.824	21	.107	.920
	Treatment	1.543	.583	29		
	Total	1.545	.686	50		
Class Boredom	Control	1.786	.747	21	1.148	.256
	Treatment	1.618	.817	29		
	Total	1.689	.785	50		
HW Enjoyment	Control	2.857	1.380	21	-1.613	.113
	Treatment	3.400	1.263	29		
	Total	3.170	1.327	50		
HW Anxiety	Control	1.809	.749	21	1.176	.245
	Treatment	1.586	.647	20		
	Total	1.680	.693	50		
HW Boredom	Control	2.395	1.210	21	1.735	.089
	Treatment	1.942	.972	29		
	Total	2.103	1.084	50		
Test Enjoyment	Control	3.016	1.331	21	-1.091	.280
	Treatment	3.425	1.046	29		
	Total	3.253	1.179	50		
Test Anxiety	Control	2.295	1.285	21	.888	.379
	Treatment	2.000	.956	29		
	Total	2.124	1.103	50		

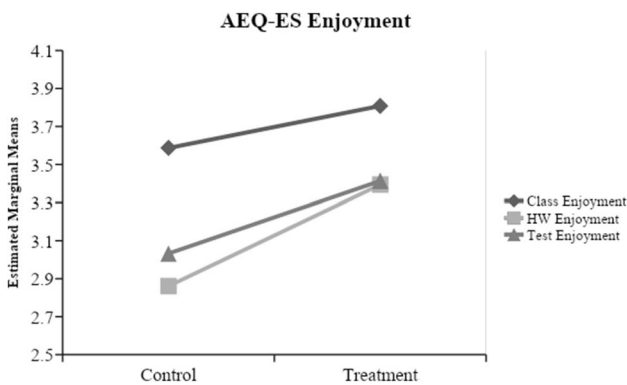


Fig. 1 Estimated marginal means of control and treatment conditions, reading enjoyment

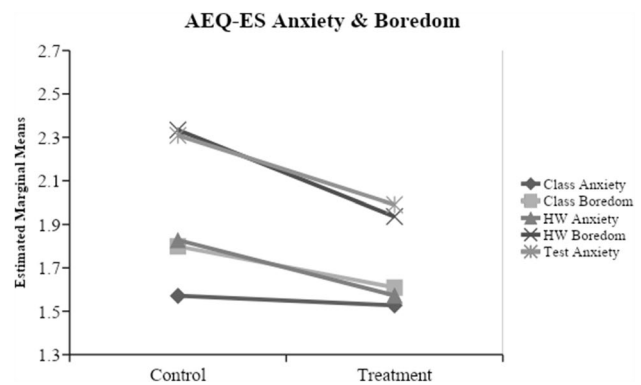


Fig. 2 Estimated marginal means of control and treatment conditions, reading anxiety and boredom

did not differ according to which condition they had been assigned to (intervention or control), or according to their GPA. However, when looking at Tests of Between-Subjects Effects (see Table 4), a significant regression equation was found, $F_{1,47} = 5.554, p < .05$. That is, *class anxiety* decreased .199 points for each point increase of one's *GPA*, $B = -.199, t(47) = -2.357, p < .05$.

Although there were no additional significant differences between treatment conditions on reading-related emotions,

a trend surfaced in the data, which is worth mentioning (see Table 5; Figs. 1 and 2). That is, as compared to participants in the control condition, participants in the treatment condition reported higher levels of *class enjoyment*, *homework enjoyment*, and *test enjoyment*; in addition, participants in the treatment condition reported lower levels of *class anxiety*, *class boredom*, *homework anxiety*, *homework boredom* and *test anxiety* than those participants in the control condition. Additionally, means of *class anxiety* were the most

similar across both conditions; *test anxiety* and *homework boredom* had higher means than *homework anxiety* and *class boredom* across both conditions.

Discussion

Core components of this self-affirmation intervention include its overall structure towards enhancing academic achievement of historically negatively stereotyped students of minoritized backgrounds, with a primary focus on group membership and self-integrity in the context of academic performance, which have continuously been demonstrated as posing a threat to academic performance stemming from the negative connotation(s) often associated with minoritized groups. (Cohen, 2006). The intervention included a written-exercise in the form of a packet, distributed to students in either the treatment or control group. The packets detailed questions in relation to students' values and their personal interpretation of the importance of these values.

Treatment group questions were geared towards their own relationship with such values, whereas the Control group questions were posed in the context of how such values may or may not be important to someone else. Students then provided their level of agreement with the values depicted in their packets in the context of question presentation depending on their randomly assigned group (treatment v. control). The use of a self-affirming intervention with minoritized students could yield successful academic outcomes, as the everyday academic experience and subsequent academic performance of minoritized students is much different than that of their non-minority peers, as minoritized students are faced with the negative stereotypical attributes of their group membership, often resulting in increased stress levels in academic settings and overall undermined performance. The intervention aimed to alter the psychological experience of minoritized students in an effort to improve academic performance, with the overarching goal of closing the racial achievement gap in the American public school system.

The seminal work of Cohen et al. (2006) has been replicated in similar studies over the last 10–15 years. Bowen et al. (2013) and Borman et al. (2016) implemented writing-based, self-affirming interventions to minoritized students in middle school (grades 6–8). Both studies noted the effects of change over time and other mediating variables that may have influenced their results with respect to the findings of Cohen et al. (2006). Such as, rates of change in grades over time (Bowen et al., 2013) and curriculum modification/reform (Borman et al., 2016). Essentially, the magnitude of a study's findings must be taken into consideration within the context of the specific educational environment studied, as well as student academic growth over time. Understanding

the mediating and moderating effects of the populations being studied can serve as helpful insight into the findings of any study utilizing a self-affirming intervention in a K-12 educational setting, as educational curricula changes rapidly overall, varies state-to-state, and no two educational environments are identical.

Following the seminal work of Cohen et al. (2006), whereby the academic achievement gap between Black and White middle school students were reduced by 40% through the use of a self-affirmation intervention, this study examined the effectiveness of such an intervention on 4th grade students' academic achievement levels. Overall, no significant difference in improvement of achievement scores was found among students who were given the intervention (i.e., treatment) as compared to students in the control condition, and the intervention was not a significant contributor to the variance found in standardized reading achievement scores or literacy GPA. Rather, pre-HMH scores (i.e., standardized reading test scores before the intervention) explained a significant amount of variance in standardized reading scores after intervention (Post-HMH) and in literacy grades (GPA). In addition, the intervention did not predict Post-HMH or GPA scores differentially across ethnic groups.

Finally, the study did not find significant influences of the intervention on reading-related emotions. These results are disappointing, as addressing academic achievement discrepancies in the earlier years was the goal of this study. Although the study did not yield significant results, it is important to recognize the notion of counterfactual comparison (Lemons et al., 2014). Counterfactual comparison of past experiments may not accurately reflect counterfactual comparisons between groups of present or future studies. No two studies and no two comparison groups are the same, especially control groups (Lemons et al., 2014). In the context of the present study, this could contribute to the differing results noted in this study in comparison to the results of Cohen et al. (2006). However, some noteworthy trends did arise.

Upon further examination, it was found that participants in the treatment condition reported higher levels of enjoyment in class, homework, and tests as related to reading as compared to their peers in the control condition; in addition, they reported lower levels of anxiety and boredom in the class and in doing homework, as well as lower test anxiety than their peers who were in the control condition. Finding significant results in these trends is of critical importance, as emotions related to academics have been found to influence motivation, performance, and physical and mental health in schools (Carmona-Halty et al., 2019; Parker & Wampler, 2006; Pekrun, 2006).

Several hypotheses were developed in attempting to understand why the intervention did not prove effective among the fourth-graders. As part of the intervention,

students were provided with a manipulation reinforcement after they had finished their written assignment, whereby the students indicated their level of agreement with phrases such as “I care about these values” (or “some people care about these values” for those in the control condition). This part of the intervention raised confusion among many students, who did not understand the concept of the Likert-scale or what they were being asked to do. This level of confusion may have dampened the self-affirmation effects that older students benefited from previous studies.

Additionally, the developmental appropriateness of the self-affirmation interventions for fourth-grade students was questioned in relation to the effectiveness of this intervention. Self-affirmation has been observed to work well with adolescents and may be an especially good fit for that developmental stage because of the strong inclination to develop social identity. Despite evidence that fourth-grade students find stereotypes to be salient in their lives and social identities (Byrd, 2012), it is quite possible that they are not developmentally ready for the effects of the self-affirmation intervention to take place.

As previously noted, students in the self-affirmation condition reported higher levels of class enjoyment, homework enjoyment, and test enjoyment than their peers in the control condition, though these differences were not statistically significant. Conversely, they reported lower levels – though not significantly so – of class anxiety, class boredom, homework anxiety, homework boredom, and test anxiety than their peers in the control condition. It was also observed that the greatest difference in means between the conditions were in the situations of test anxiety and homework boredom; while experiencing these situational feelings, it is conceivable that students are susceptible to threatening thoughts related to self-esteem and stereotypes. Stereotype threat is theorized to correlate with high levels of anxiety and evaluation apprehension (Smith, 2004), which could be the case here, had results been found to be significant.

Implications for Future Research and Practice

Academic underperformance of stereotyped minoritized students can be explained as a recursive process (Cohen et al., 2009), underscoring the importance of psychosocial interventions that start in elementary school. In fact, such interventions should be implemented even earlier, focusing on the school readiness gap by promoting positive parenting and valuing achievement at home (Social Equity Theory; McKown, 2013). The relationship of mental health, specifically anxiety and depression, academic, and physical health outcomes related to bias, stigma and stereotype threat is substantial (Sukhera et al., 2019). There is a need for mind-body health interventions such as, positive self-affirmations, to improve social emotional

functioning leading to strengthened self-efficacy and the motivation (Gillen-O’Neel et al., 2011) and ability to take action when faced with barriers such as stigma, bias, and threat. Medically underserved and those who are educationally and economically disadvantaged, in particular, are burdened by these obstacles.

The role of the school psychologist and other school-based mental health professionals is also important to discuss. Their role includes focus on these important areas. There are a myriad of mind-body health treatments that can be applied at the individual and group level (e.g., classroom and/or school) that effectively combat stereotype threats, including academic, as well as various biases and stigma. The goal being to increase relationships and belongingness in groups and respect for individual differences (Rosenthal & Crisp, 2006).

While no assertions can be made about the efficacy of this particular intervention for elementary school students on the basis of this study, students in the self-affirmation condition reported enjoying class, homework, and tests more than students in the control condition. Thus, educators might be disposed to implement a short self-affirmation activity to bolster student engagement. This is in line with extant literature that explains changing the subjective experience about one’s environment are sound practices to be made in schools in attempts to increase achievement levels (Yeager et al., 2013). Furthermore, interventions such as the one used here are quite readily available for use on a class- and school-wide level.

Future investigations on self-affirmation intervention with younger students should take care to avoid elements that could confuse students by introducing unnecessary complexity, as the Likert scale questions appeared to do in this study during the manipulation reinforcement. Researchers conducting similar studies would be advised to simplify and/or restructure such a procedure or forego it completely.

According to regulatory fit, stereotype threat causes individuals to avoid failure and thus performance is increased through the adoption of the avoidance goal, whereas individuals who adopt a performance approach goal while under stereotyped conditions will see reduced performance (Chalabaev et al., 2012). Researchers interested in exploring emotional functioning as related to classroom performance might investigate the connection between the emotions assessed here (i.e., enjoyment, anxiety, and boredom) in addition to self-efficacy as related to performance avoidance goals. Given the significant disconnect that still exists between psychology and education (Weinstein, 2002), school-based studies like these are needed to close this particular gap before closing the achievement gap becomes a reality.

Data Availability The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

This material is the authors' own original work, which has not been previously published elsewhere. The paper is not currently being considered for publication elsewhere. The paper reflects the authors' own research and analysis in a truthful and complete manner.

Ethical Statement This study was approved by the University of Connecticut Institutional Review Board.

Informed Consent Informed consent was obtained from all individual participants involved in the study.

Conflict of Interest to Disclose The authors confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

References

- Adams, G., Garcia, D. M., Purdie-Vaughns, C., & Steele, C. M. (2006). The detrimental effects of a suggestion of sexism in an instruction situation. *Journal of Experimental Social Psychology, 42*, 602–615. <https://doi.org/10.1016/j.jesp.2005.10.004>
- Ambady, N., Shih, M., Kim, A., & Pittinsky, T. L. (2001). Stereotype susceptibility in children: Effects of identity activation on quantitative performance. *Psychological Science, 12*, 385–390. <https://doi.org/10.1111/1467-9280.00371>
- Appel, M., & Kronberger, N. (2012). Stereotypes and the achievement gap: Stereotype threat prior to test taking. *Educational Psychology Review, 24*, 609–635. <https://doi.org/10.1007/s10648-012-9200-4>
- Appel, M., Kronberger, N., & Aronson, J. (2011). Stereotype threat impedes ability building: Effects on test preparation among women in science and technology. *European Journal of Social Psychology, 41*, 904–913. <https://doi.org/10.1002/ejsp.835>
- Aronson, J. (2004). The threat of stereotype. *Educational Leadership, 14*–19.
- Aronson, J. (2005). Stereotype and the fragility of human competence, motivation, and self-concept. In C. Dweck & E. Elliot, *Handbook of competence and motivation*. Guilford Press.
- Bacharach, V. R., Baumeister, A. A., & Furr, R. M. (2003). Racial and gender science achievement gaps in secondary education. *The Journal of Genetic Psychology, 164*(1), 115–126.
- Bancroft, A., Bratter, J., & Rowley, K. (2017). Affirmation effects on math scores: The importance of high school track. *Social Science Research, 64*, 319–333. <https://doi.org/10.1016/j.ssresearch.2016.10.001>
- Binning, K. R., Cook, J. E., Greenaway, V. P., Garcia, J., Apfel, N., Sherman, D. K., & Cohen, G. L. (2021). Securing self-integrity over time: Self-affirmation disrupts a negative cycle between psychological threat and academic performance. *Journal of Social Issues, 77*(3), 801–823. <https://doi.org/10.1111/josi.12461>
- Borman, G. D., Grigg, J., & Hanselman, P. (2016). An effort to close achievement gaps at scale through self-affirmation. *Educational Evaluation and Policy Analysis, 38*(1), 21–42. <https://doi.org/10.3102/0162373715581709>
- Bowman, B. T., Comer, J. P., & Johns, D. J. (2018). Addressing the African American achievement gap: Three leading educators issue a call to action. *YC Young Children, 73*(2), 14–23.
- Byrd, C. M. (2012). The measurement of racial/ethnic identity in children: A critical review. *Journal of Black Psychology, 38*(1), 3–31. <https://doi.org/10.1177/0095798410397544>
- Carmona-Halty, M., Salanova, M., Llorens, S., & Schaufeli, W. B. (2019). How psychological capital mediates between study-related positive emotions and academic performance. *Journal of Happiness Studies, 20*(2), 605–617. <https://doi.org/10.1007/s10902-018-9963-5>
- Chalabaev, A., Major, B., Sarrazin, P., & Cury, F. (2012). When avoiding failure improves performance: Stereotype threat and the impact of performance goals. *Motivation and Emotion, 36*, 130–142. <https://doi.org/10.1007/s11031-011-9241-x>
- Cheryan, S., Plaut, V. C., Davies, P., & Steele, C. M. (2009). Ambient belonging: How stereotypical environments impact gender participation in computer science. *Journal of Personality and Social Psychology, 97*, 1045–1060. <https://doi.org/10.1037/a0016239>
- Cohen, G. L., & Sherman, D. K. (2014). The psychology of change: Self-affirmation and social psychological intervention. *Annual Review of Psychology, 65*(1), 333–371. <https://doi.org/10.1146/annurev-psych-010213-115137>
- Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science, 313*(5791), 1307–1310.
- Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., & Brzuscoski, P. (2009). Recursive processes in self-affirmation: Intervening to close the minority achievement gap. *Science, 324*, 400–403.
- Cook, J. E., Purdie-Vaughns, V., Garcia, J., & Cohen, G. L. (2012). Chronic threat and contingent belonging: Protective benefits of values affirmation on identity development. *Journal of Personality and Social Psychology, 102*(3), 479, 96.
- Corenblum, B., & Armstrong, H. D. (2012). Racial-ethnic identity development in children in a racial-ethnic minority group. *Canadian Journal of Behavioural Science, 44*(2), 124–137.
- Critcher, C. R., & Dunning, D. (2015). Self-affirmations provide a broader perspective on self-threat. *Personality and Social Psychology Bulletin, 41*(1), 3–18. <https://doi.org/10.1177/0146167214554956>
- Depaoli, J. L., Balfanz, R., Atwell, M. N., & Bridgeland, J. (2018). Building a Grad Nation: Progress and Challenge in Raising High School Graduation Rates. Annual Update 2018. *Civic Enterprises*. <https://files.eric.ed.gov/fulltext/ED585524.pdf>
- Desert, M., Preaux, M., & Jund, R. (2009). So young and already victims of stereotype threat: Socio-economic status and performance of 6 to 9 years old children on Raven's progressive matrices. *European Journal of Psychology of Education, 24*, 207–218. <https://doi.org/10.1007/BF03173012>
- Fryer, R. G., & Levitt, S. D. (2004). Understanding the black-white test score gap in the first two years of school. *The Review of Economics and Statistics, 86*(2), 447–464. <https://doi.org/10.1162/003465304323031049>
- Gillen-O'Neel, C., Ruble, D. N., & Fuligni, A. J. (2011). Ethnic stigma, academic anxiety, and intrinsic motivation in middle childhood. *Child Development, 82*(5), 1470–1485. <https://doi.org/10.1111/j.1467-8624.2011.01621.x>
- Hanushek, E. A., Peterson, P. E., Talpey, L. M., & Woessmann, L. (2019). The achievement gap fails to close. *Education Next, 19*(3), 8–17.
- Hartley, B. L., & Sutton, R. M. (2013). A stereotype threat account of boys' academic underachievement. *Child Development, 85*(5), 1716–1733. <https://doi.org/10.1111/cdev.12079>
- Howard, S., & Borgella, A. (2018). "Sinking" or sinking?: Identity salience and shifts in black women's athletic performance.

- Psychology of Sport and Exercise*, 39, 179–183. <https://doi.org/10.1016/j.psychsport.2018.08.016>
- Howe, M. J. A., Davidson, J. W., & Sloboda, J. A. (1998). Innate talents: Reality or myth? *Behavioral and Brain Sciences*, 21, 399–442. <https://doi.org/10.1017/S0140525X9800123X>
- Jordan, A. H., & Lovett, B. J. (2007). Stereotype threat and test performance: A primer for school psychologists. *Journal of School Psychology*, 45, 45–59. <https://doi.org/10.1016/j.jsp.2006.09.003>
- Juvonen, J. (2007). Reforming middle schools: Focus on continuity, social connectedness, and engagement. *Educational Psychologist*, 42, 197–208. <https://doi.org/10.1080/00461520701621046>
- Kellow, J. T., & Jones, B. D. (2008). The effects of stereotypes on the achievement gap: Reexamining the academic performance of African American high school students. *Journal of Black Psychology*, 34(1), 94–120. <https://doi.org/10.1177/0095798407310537>
- Klein, W. M. P., Harris, P. R., Ferrer, R. A., & Zajac, L. E. (2011). Feelings of vulnerability in response to threatening messages: Effects of self-affirmation. *Journal of Experimental Social Psychology*, 47, 1237–1242. <https://doi.org/10.1016/j.jesp.2011.05.005>
- Korpershoek, H., Canrinus, E. T., Fokkens-Bruinsma, M., & de Boer, H. (2020). The relationships between school belonging and students' motivational, social-emotional, behavioural, and academic outcomes in secondary education: A meta-analytic review. *Research Papers in Education*, 35(6), 641–680. <https://doi.org/10.1080/02671522.2019.1615116>
- Lemons, C. J., Fuchs, D., Gilbert, J. K., & Fuchs, L. S. (2014). Evidence-based practices in a changing world. *Educational Researcher*, 43(5), 242–252. <https://doi.org/10.3102/0013189x14539189>
- Lichtenfeld, S., Pekrun, R., Stupnisky, R. H., Reiss, K., & Murayama, K. (2012). Measuring students' emotions in the early years: The achievement emotions questionnaire-elementary school (AEQ-ES). *Learning and Individual Differences*, 22(2), 190–201. <https://doi.org/10.1016/j.jesp.2011.05.005>
- Liu, S., Liu, P., Wang, M., & Zhang, B. (2021). Effectiveness of stereotype threat interventions: A meta-analytic review. *Journal of Applied Psychology*, 106(6), 921. <https://psycnet.apa.org/doi/10.1037/apl0000770>
- Mark, D. L. H. (2013). Academic achievement gap or gap of opportunities? *Urban Education*, 48(2), 335–343. <https://doi.org/10.1177/0042085913476936>
- Martens, A., Johns, M., Greenberg, J., & Schimel, J. (2006). Combating stereotype threat: The effect of self-affirmation on women's intellectual performance. *Journal of Experimental Social Psychology*, 42, 236–243. <https://doi.org/10.1016/j.jesp.2005.04.010>
- Marx, D. M., Ko, S. J., & Friedman, R. A. (2009). The “Obama effect”: How a salient role model reduces race-based performance differences. *Journal of Experimental Social Psychology*. <https://doi.org/10.1016/j.jesp.2009.03.012>
- McKown, C. (2013). Social equity theory and racial-ethnic achievement gaps. *Child Development*, 84(4), 1120–1136. <https://doi.org/10.1111/cdev.12033>
- Mello, Z. R., Mallett, R. K., Andretta, J. R., & Worrell, F. C. (2012). Stereotype threat and school belonging in adolescents from diverse racial/ethnic backgrounds. *Journal of At-Risk Issues*, 17(1), 9–14.
- Milner, R. (2012). Beyond a test score: Explaining opportunity gaps in educational practice. *Journal of Black Studies*, 43(6), 693–718. <https://doi.org/10.1177/0021934712442539>
- Murphy, M. C., Gopalan, M., Carter, E. R., Emerson, K. T., Bottoms, B. L., & Walton, G. M. (2020). A customized belonging intervention improves retention of socially disadvantaged students at a broad-access university. *Science Advances*, 6(29), [eaba4677]. <https://doi.org/10.1126/sciadv.aba4677>
- National Center for Education Statistics. (2009). *Achievement gaps: How black and white students in public schools perform in mathematics and reading on the national assessment of educational progress*. National Center for Education Statistics. <https://nces.ed.gov/nationsreportcard/pdf/studies/2009455.pdf>
- National Center for Education Statistics. (2011). *Achievement gaps: How Hispanic and white students in public schools perform in mathematics and reading on the national assessment of educational progress*. National Center for Education Statistics. <https://nces.ed.gov/nationsreportcard/pdf/studies/2011459.pdf>
- National Center for Education Statistics. (2019). *Achievement gaps dashboard*. The Nation's Report Card. https://www.nationsreportcard.gov/dashboards/achievement_gaps.aspx
- Nelson, S. K., Fuller, J. A. K., Choi, I., & Lyubomirsky, S. (2014). Beyond self-protection: Self-affirmation benefits hedonic and eudemonic well-being. *Personality and Social Psychology Bulletin*, 40(8), 998–1011. <https://doi.org/10.1177/0146167214533389>
- Oates, G. L. S. C. (2009). An empirical test of five prominent explanations for the black-white academic performance gap. *Social Psychological Education*, 12, 415–441. <https://doi.org/10.1007/s1218-009-9091-5>
- Olszewski-Kubilius, P., & Corwith, S. (2018). Poverty, academic achievement, and giftedness: A literature review. *Gifted Child Quarterly*, 62(1), 37–55. <https://doi.org/10.1177/0016986217738015>
- Owens, J., & Massey, D. S. (2011). Stereotype threat and college academic performance: A latent variables approach. *Social Science Research*, 40, 150–166. <https://doi.org/10.1016/j.ssresearch.2010.09.010>
- Parker, T. S., & Wampler, K. S. (2006). Changing emotion: The use of therapeutic storytelling. *Journal of Marital and Family Therapy*, 32(2), 155–166. <https://doi.org/10.1111/j.1752-0606.2006.tb01597.x>
- Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, 18, 315–341. <https://doi.org/10.1007/s10648-006-9029-9>
- Riciputi, S., & Erdal, K. (2017). The effect of stereotype threat on student-athlete math performance. *Psychology of Sport and Exercise*, 32, 54–57. <https://doi.org/10.1016/j.psychsport.2017.06.003>
- Rosenthal, H. E. S., & Crisp, R. J. (2006). Reducing stereotype threat by blurring intergroup boundaries. *Personality & Social Psychology Bulletin*, 32(4), 501–511. <https://doi.org/10.1177/0146167205281009>
- Rossi-Arnaud, C., Spataro, P., & Geraci, L. (2018). Effects of stereotype threat and prior task success on older adults' eyewitness memory. *Journal of Applied Research in Memory and Cognition*, 7(3), 422–431. <https://doi.org/10.1016/j.jarmac.2018.02.001>
- Rydell, R. J., Rydell, M. T., & Boucher, K. L. (2010). The effect of negative performance stereotypes on learning. *Journal of Personality and Social Psychology*, 99(6), 883–896. <https://doi.org/10.1037/a0021139>
- Schmader, T., & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology*, 85(3), 440–452. <https://doi.org/10.1037/0022-3514.85.3.440>
- Sherman, D. K., Cohen, G. L., Nelson, L. D., Nussbaum, A. D., Bunyan, D. P., & Garcia, J. (2009). Affirmed yet unaware: Exploring the role of awareness in the process of self-affirmation. *Journal of Personality and Social Psychology*, 97(5), 745–764. <https://doi.org/10.1037/a0015451>
- Smith, J. L. (2004). Understanding the process of stereotype threat: A review of meditational variables and new performance goal directions. *Educational Psychology Review*, 16, 177–206. <https://doi.org/10.1023/B:EDPR.0000034020.20317.89>
- Spencer, S. J., Logel, C., & Davies, P. G. (2016). Stereotype threat. *Annual Review of Psychology*, 67, 415–437. <https://doi.org/10.1146/annurev-psych-073115-103235>

- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*(5), 797–811. <https://doi.org/10.1037/0022-3514.69.5.797>
- Steele, C. M., Spencer, S. J., & Aronson, J. (2002). Contending with group image: The psychology of stereotype and social identity threat. *Advances in Experimental Social Psychology*, *34*, 379–440. [https://doi.org/10.1016/S0065-2601\(02\)80009-0](https://doi.org/10.1016/S0065-2601(02)80009-0)
- Stephens, N. M., Markus, H. R., & Fryberg, S. A. (2012). Social class disparities in health and education: Reducing inequality by applying a sociocultural self model of behavior. *Psychological Review*, *119*(4), 723–744. <https://doi.org/10.1037/a0029028>
- Sukhera, J., Wodzinski, M., Milne, A., Teunissen, P. W., Lingard, L., & Watling, C. (2019). Implicit bias and the feedback paradox: Exploring how health professionals engage with feedback while questioning its credibility. *Academic Medicine*, *94*(8), 1204–1210. <https://doi.org/10.1097/ACM.0000000000002782>
- Taylor, V. J., & Walton, G. M. (2011). Stereotype threat undermines academic learning. *Personality and Social Psychology Bulletin*, *37*(8), 1055–1067. <https://doi.org/10.1177/0146167211406506>
- Thoman, D. B., Smith, J. L., Brown, E. R., Chase, J., & Lee, J. Y. K. (2013). Beyond performance: A motivational experiences model of stereotype threat. *Educational Psychology Review*, *25*, 211–243. <https://doi.org/10.1007/s10648-013-9219-1>
- Tomasetto, C., Alparone, F. R., & Cadinu, M. (2011). Girls' math performance under stereotype threat: The moderating role of mothers' gender stereotypes. *Developmental Psychology*, *47*, 943–949.
- Umaña-Taylor, A. J., Yazedjian, A., & Bamaca-Gomez, M. (2009). Developing the ethnic identity scale using Eriksonian and social identity perspectives. *Identity: An International Journal of Theory and Research*, *4*(1), 9–38. [10.1207/S1532706XID0401_2](https://doi.org/10.1207/S1532706XID0401_2)
- Umaña-Taylor, A. J., Quintana, S. M., Lee, R. M., Cross, W. E., Rivas-Drake, D., Schwartz, S. J., et al. (2014). Ethnic and racial identity during adolescence and into young adulthood: An integrated conceptualization. *Child Development*, *85*, 21–39. <https://doi.org/10.1111/cdev.12196>
- Vermeulen, L., Castellar, E. N., Janssen, D., Calvi, L., & Van Looy, J. (2016). Playing under threat. Examining stereotype threat in female game players. *Computers in Human Behavior*, *57*, 377–387. <https://doi.org/10.1016/j.chb.2015.12.042>
- Voight, A. M., Geller, J. D., & Nation, M. (2014). Contextualizing the “behavior gap”: Student prosocial behavior and racial composition in urban middle schools. *Journal of Early Adolescence*, *34*(2), 157–177. <https://doi.org/10.1177/0272431613482043>
- Walton, G. M., & Cohen, G. L. (2007). A question of belonging: Race, social fit, and achievement. *Journal of Personality and Social Psychology*, *92*, 82–96. <https://doi.org/10.1037/0022-3514.92.1.82>
- Walton, G. M., Cohen, G. L., Cwir, D., & Spencer, S. J. (2012). Mere belonging: The power of social connections. *Journal of Personality and Social Psychology*, *102*(3), 513–532. <https://doi.org/10.1037/a0025731>
- Weinstein, R. S. (2002). *Reaching higher: The power of expectations in schooling*. Harvard University Press.
- Wout, D., Danso, H., Jackson, J., & Spencer, S. (2008). The many faces of stereotype threat: Group- and self-threat. *Journal of Experimental Social Psychology*, *44*, 792–799. <https://doi.org/10.1016/j.jesp.2007.07.005>
- Yeager, D., Walton, G., & Cohen, G. L. (2013). Addressing achievement gaps with psychological interventions. *Phi Delta Kappan*, *94*(5), 62–65. <https://doi.org/10.1177/003172171309400514>
- Yeung, W. J., & Conley, D. (2008). Black-white achievement gap and family wealth. *Child Development*, *79*(2), 303–324. <https://doi.org/10.1111/j.1467-8624.2007.01127.x>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.