Scientific Research Publishing

Open Journal of Social Sciences, 2026, 14(4), 239-257
https://www.scirp.org/journal/jss
ISSN Online: 2327-5960
ISSN Print: 2327-5952

Does Self-Affirmation Improve Educational Outcome of Students? A Structured Review of Self-Affirmation Interventions

Fuqing Han

School of Education, Durham University, Durham, United Kingdom
Email: fpjf80@durham.ac.uk

How to cite this paper: Han, F. Q. (2026). Does Self-Affirmation Improve Educational Outcome of Students? A Structured Review of Self-Affirmation Interventions. *Open Journal of Social Sciences, 14,* 239-257.
https://doi.org/10.4236/jss.2026.144014

Received: March 13, 2026
Accepted: April 13, 2026
Published: April 16, 2026

Abstract

Educational achievement gaps persist globally, with some ethnic minority and socioeconomically disadvantaged students consistently underperforming. Self-affirmation interventions, brief value-reflection exercises, have shown promise for reducing identity threats, yet recent replications yield mixed results, necessitating systematic evaluation of when and how these interventions might work. This structured review evaluates self-affirmation interventions for adolescents aged 11 - 15, addressing: 1) overall effectiveness; 2) underlying mechanisms; 3) moderating factors; 4) optimal conditions; and 5) implications for practice. Following PRISMA 2020 guidelines, we systematically searched multiple databases for studies published 1990-2025, focusing on interventions for students aged 11 - 15 with educational or psychological outcomes. The synthesis of eighteen studies involving 81 to 4002 participants across six countries revealed that twelve studies (67%) reported positive effects on academic performance, while four found null effects and two showed negative effects in specific subgroups. Notably, effectiveness was strongly context-dependent: interventions succeeded when moderate stereotype threat existed alongside supportive environments, multi-session implementation, and among moderately performing students. The evidence supports decoupling mechanisms whereby affirmation disrupts links between poor performance and diminished belonging, with effects proving stronger for academic outcomes and belonging than psychological distress. These findings indicate that self-affirmation interventions are context-sensitive tools requiring alignment of meaningful threat, environmental support, quality implementation, and student receptivity. Importantly, effectiveness cannot be assumed from delivery alone, and psychological interventions must complement, not replace, structural reforms. Future research should prioritize mediation testing, cultural adaptation, and implementation fidelity to advance both theoretical understanding and practical application.

Keywords

Self-Affirmation, Educational Interventions, Achievement Gaps, Stereotype Threat, Students, Structured Review

1. Introduction

Educational achievement gaps remain a persistent global challenge, with minority and low-socioeconomic status students consistently underperforming compared to their peers (Hadden et al., 2020). These disparities perpetuate social inequalities and limit opportunities for disadvantaged groups (Sherman et al., 2013). Brief psychological interventions offer cost-effective approaches to address these gaps, with self-affirmation interventions emerging as particularly promising tools for reducing identity threats in educational settings (Cohen & Sherman, 2014). Understanding the psychological mechanisms underlying these achievement gaps is crucial for developing effective interventions.

Psychological interventions offer cost-effective approaches to address these gaps, with self-affirmation interventions emerging as particularly promising tools for reducing identity threats in educational settings (Cohen & Sherman, 2014). Accordingly, this study focuses on self-affirmation interventions as a key mechanism for understanding and addressing educational inequalities.

Self-affirmation interventions are grounded in two complementary theories: stereotype threat theory (Steele & Aronson, 1995) and self-affirmation theory (Steele, 1988). The basic premise is that brief value-reflection exercises buffer psychological threats to academic identity. Further theoretical explanations are provided below.

Translating these theoretical principles into practice, researchers have developed simple, scalable classroom-based interventions. Self-affirmation interventions typically involve brief (15 - 20 minute) writing exercises where students select and reflect on personally important values such as family, friendships, or creative expression (Cohen et al., 2006). These exercises are administered multiple times throughout an academic term and require minimal resources, making them highly scalable (Escobar-Soler et al., 2024). Control participants write about values least important to them, maintaining similar task demands without self-affirming content.

Empirical evidence reveals that self-affirmation interventions affect both psychological processes and academic outcomes. At the cognitive level, affirmed students demonstrate reduced defensive processing of threatening feedback, increased receptiveness to constructive criticism, and more abstract (versus concrete) construal of academic challenges (Sherman et al., 2013). Psychologically, self-affirmation reduces evaluation anxiety, decreases cortisol responses to academic stress, and promotes a growth-oriented mindset (Creswell et al., 2013). These psychological shifts translate into behavioral changes: affirmed students show increased

help-seeking behavior, greater engagement with challenging material, and improved belonging perceptions (Cook et al., 2012).

Cohen et al. (2006, 2009) found that these cumulative psychological and behavioral changes reduced racial achievement gaps by 40%, with effects persisting through graduation. However, recent replications have yielded mixed results. Hanselman et al. (2017) found no evidence of benefits when replicating the intervention, suggesting that the psychological mechanisms may be more context-dependent than initially theorized. A meta-analysis by Escobar-Soler et al. (2024) found a small overall effect (d = 0.41) with substantial heterogeneity, highlighting the need to better understand when and how these cognitive processes are activated.

This structured review addresses this need by systematically evaluating self-affirmation intervention research published between 1990 and 2025. Specifically, this review addresses five primary research questions:

1) What is the overall effect of self-affirmation interventions on student educational outcomes?

2) Through what cognitive and psychological mechanisms do self-affirmation interventions operate?

3) Which contextual, individual, and implementation factors moderate intervention effectiveness?

4) For whom and under what conditions are these cognitive benefits most pronounced?

5) What implications does the current evidence hold for educational practice and policy?

2. Methods

This section outlines the methodological approach adopted in this study, including the search strategy, eligibility criteria, screening process, data extraction, and quality appraisal procedures.

2.1. Methodological Approach

A systematic literature search was conducted to identify self-affirmation intervention studies published from 1990 to 2025.

Multiple electronic databases were searched via EBSCOhost (including APA PsycInfo, Education Research Complete, and ERIC), ProQuest (for grey literature including theses and dissertations), Scopus, and Web of Science.

The search strategy employed Boolean operators combining three core components: sample characteristics (e.g., "students", "adolescents"), intervention type (e.g., "self-affirmation"), and outcomes (e.g., "academic achievement", "school belonging", "stereotype threat"). The specific search string was:

("self-affirmation intervention*" OR "self-affirmation intervention*" OR "values affirmation") AND ("adolescent*" OR "teenager*" OR "youth" OR "student*") AND ("academic achieve*" OR "attain*" OR "educational outcome*" OR "learning outcome*" OR "academic performance*" OR "test*" OR "score*" OR "grade*"

OR "GPA*" OR "school experience*" OR "peer relate*" OR "belonging*" OR "engagement*")

Title and abstract screening was conducted by one reviewer. To enhance reliability, a second reviewer independently checked a subset of the records.

2.2. Eligibility Criteria and Study Selection

Studies were included if they met predefined eligibility criteria (Table 1).

Table 1. Inclusion and exclusion criteria.

Criterion	Inclusion	Exclusion
Study design	Experimental or quasi-experimental designs with control groups	Observational studies, qualitative studies, case studies
Publication period	1990-2025	Studies published before 1990
Participants	Students aged 11 - 15	Clinical populations, non-student samples, elementary school students
Setting	Educational contexts (schools, universities) Non-educational settings (e.g., clinical, workplace)	Educational contexts (schools, universities) Non-educational settings (e.g., clinical, workplace)
Intervention type	Self-affirmation interventions involving value-affirmation writing exercises	Interventions combining self-affirmation with other major psychological interventions Non-writing-based self-affirmation Self-affirmation without clear manipulation check
Outcomes	Academic performance (grades, test scores, GPA) Psychological well-being (anxiety, stress, self-esteem, belonging) Educational engagement (participation, persistence, attendance)	Studies without quantitative outcome data or with only qualitative outcomes
Language	English	Non-English publications

2.3. The Process of Screening

The screening process followed a structured, multi-step approach in accordance with PRISMA 2020 guidelines. All references retrieved from the database searches were first imported into Excel, where duplicates were identified and removed manually or using Excel's built-in filtering functions. After de-duplication, a total of 3053 unique records remained for screening. Supplementary searches were conducted to ensure comprehensive coverage of relevant literature, particularly for studies not indexed in the primary databases. After de-duplication, a total of 9780 unique records remained for screening.

Title and abstract screening was conducted by the first reviewer, using a pre-designed Excel spreadsheet that included key fields such as article title, author, publication year, population age range, intervention type, outcomes measured, and initial inclusion/exclusion decisions. Studies clearly irrelevant to the topic (e.g., non-educational settings, adult samples, non-intervention studies) were excluded at this stage.

For the full-text screening phase, potentially eligible studies were retrieved in

full and re-assessed against the inclusion criteria. Each study was coded in Excel with updated columns for full-text eligibility decision, reasons for exclusion (if applicable), and notes on methodology or theoretical alignment (e.g., use of values-affirmation procedures). Ambiguities or borderline cases were discussed and resolved in consultation with a second reviewer.

In total, 19 studies met all inclusion criteria and were included in the final synthesis. A PRISMA flow diagram summarizes the entire screening and selection process. See **Figure 1**. Note on Wu (2024) inclusion: This review prioritizes studies with participants aged 11 - 15 years. One study (Wu, 2024) was included despite participants having a mean age of 17.4 years (SD = 2.31) because: 1) it examines deaf students, a marginalized and understudied population in affirmation research; 2) Wu (2024) notes that most deaf students were 2 - 4 years older than the intended age for their grade (p. 13), reflecting typical educational development delays in deaf school contexts; and 3) the study employs rigorous RCT methodology (N = 269) in a low-resource international setting (Nepal), providing important evidence on intervention effectiveness across diverse populations. This targeted inclusion reflects recognition that educational developmental patterns should inform decisions about age-eligibility for studies with understudied populations (see **Figure 1**).

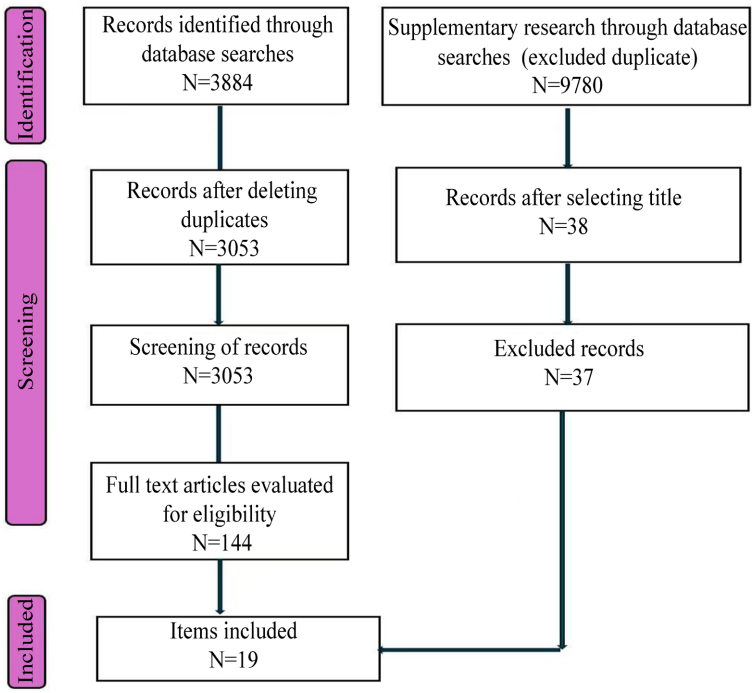

Figure 1. Prisma.

2.4. Quality Assessment

To ensure the credibility and interpretability of findings, this review incorporated a structured quality appraisal process. Rather than excluding studies based on design alone, the goal was to critically evaluate the strength of evidence and its im-

plications for synthesis (Ahn & Kang, 2018). In line with the previous structured review forming part of this doctoral research, the same "sieve" quality appraisal framework (Gorard, 2024) was employed to ensure methodological consistency and comparability (see Table 2).

Table 2. Gorard's sieve method quality assessment criteria.

Design	Scale	Missing data	Measurement quality	Rating
Well-matched experimental design	Sufficient cases per group	Negligible dropout, no impact	Consistently applied, externally validated, generally reliable	4
Well-matched experimental design	Moderate sample size per group	Some imbalance or minor attrition	Well-structured and validated, though with minor inconsist-encies	3
Roughly matched comparison	Small sample size per group	Initial imbalance or moderate loss	Lacking consistency or impartiality, with potential for significant inaccuracies	2
Poor or non-equivalent control	Very small sample size	Initial imbalance or moderate loss	Poorly constructed indicators, substantial error margins, or excessive outcome variables	1
No comparator or unclear	Undefined	Attrition unreported or excessive	Highly unreliable or invalid assessment tools	0

This tool examines a range of validity threats, such as weak research design, insufficient sample size, missing data, or unreliable outcome measures. Importantly, evaluation was based solely on methodological criteria rather than journal quality or author profile (see Table 3 for criteria). Two reviewers independently rated each study, and disagreements were resolved through discussion and reference to the predefined scoring framework, thereby ensuring inter-rater reliability.

Table 3 presents the quality appraisal of the included studies using Gorard's sieve method, which evaluates methodological strength across five dimensions: design, scale, dropout, validity, and fidelity. Each dimension was scored on a three-point scale, with higher scores indicating stronger methodological quality.

Overall, the majority of studies were rated as medium to high quality. Most studies achieved the highest score (3) on research design, reflecting the predominance of randomised controlled trials in the sample. However, lower scores were more commonly observed in the dimensions of scale, dropout, validity, and fidelity, suggesting limitations related to sample size, attrition, and implementation consistency.

Only a small number of studies (e.g., Borman et al., 2018; Goyer et al., 2017; Hoffman et al., 2019; Yan et al., 2023) achieved consistently high scores across all dimensions, indicating stronger overall methodological robustness. In contrast, several studies received lower ratings in multiple domains, highlighting variability in evidence quality across the included literature.

Table 3. Quality appraisal of included studies using Gorard's sieve method.

Article	Design	Scale	Dropout	Validity	Fidelity	Rate
Cohen et al. (2006)	3	2	2	2	2	2
Cohen et al. (2009)	3	2	2	2	2	2
Sherman et al. (2013)	3	2	2	2	2	2
Dee (2015)	3	3	3	3	2	3
Borman et al. (2018)	3	3	3	3	3	3
Goyer et al. (2017)	3	3	3	3	3	3
Covarrubias et al. (2016)	3	2	2	2	2	2
Hoffman et al. (2019)	3	3	3	3	3	3
Hernandez et al. (2017)	3	2	2	2	2	2
Lokhande & Müller (2019)	3	3	3	3	3	3
Wu (2024)	3	2	2	2	2	2
Protzko & Aronson (2016)	2	3	2	2	2	2
Hadden et al. (2020)	3	3	3	3	3	3
Bratter et al. (2016)	3	2	2	2	2	2
Cook et al. (2012)	3	2	2	2	2	2
Yan et al. (2023)	3	3	3	3	3	3
Binning et al. (2019)	3	2	2	2	2	2
See et al. (2022)	3	3	3	3	3	3
Yang et al., (2023)	3	2	2	2	2	2

These results suggest that, while the evidence base is generally of moderate quality, caution is warranted when interpreting findings, particularly for studies with lower scores in validity and fidelity.

3. Results

This section presents the findings of the review. It first describes the characteristics of the included studies. The main findings are then synthesised according to key outcome domains, including academic performance and school experience.

3.1. Overall Self-Affirmation Interventions

Most studies included academic performance or school experience outcomes in their measurements. A small number of studies documented positive effects on academic performance (Cohen et al., 2006, 2009; Sherman et al., 2013; Borman et al., 2018; Goyer et al., 2017; Hernandez et al., 2017; Cook et al., 2012; Binning et al., 2019; See et al., 2022; Yang et al., 2023), with effect sizes ranging from GPA

improvements of 0.24 - 0.34 points to reductions in achievement gaps (45% - 50%) and behavioral improvements (69% lower discipline rates). One study (Dee, 2015) reported mixed effects, with positive outcomes for Hispanic 8th graders but negative effects for female 8th graders.

Three studies reported null or negative findings: Hoffman et al. (2019) found that single-session interventions were insufficient; Bratter et al. (2016) reported no clear evidence of benefit with negative effects for some groups; and Covarrubias et al. (2016) found family affirmation substantially more effective than self-affirmation alone. Three additional studies (Wu, 2024; Protzko & Aronson, 2016; Hadden et al., 2020) reported primarily baseline group comparisons or had incomplete outcome data. One study (Yan et al., 2023) found effects on well-being and mental health but examined psychological rather than academic outcomes. Overall, approximately two-thirds of studies (67%, n = 12) demonstrated positive or context-dependent effects.

The sample sizes varied substantially across studies. The smallest sample comprised 82 participants, while the largest initially recruited 4002 participants, of whom 2234 were retained for final analysis after quality screening. Most participants were students aged 11 - 15 years, aligning with the age range specified in the review criteria; however, one study (Wu, 2024) included participants with a higher mean age (17.4 years). This study was retained because the sample included adolescents within the target age range (11 - 15) and focused on deaf students, for whom chronological age may not directly correspond to educational stage due to differences in educational progression. Including this study also allows the review to capture evidence from an underrepresented and marginalized population, consistent with the aims of this review.

In terms of research design, 18 studies adopted randomized controlled trial (RCT) designs, while one study (Hernandez et al., 2017) employed a quasi-experimental design. All studies implemented single-blind procedures for participants, meaning that students were unaware of the study hypothesis and group assignment.

Across studies, the form of intervention was consistent: all employed written value-affirmation tasks. However, the frequency of administration varied. Few studies implemented a single session, seven conducted two or three sessions, and four or more conducted four or five sessions. Each single session typically lasted between 10 and 20 minutes, though one study used a 30-minute session.

The target populations reflected a gradual evolution in research focus. Earlier studies concentrated primarily on ethnic minority students in the United States, particularly African American and Latino populations. Subsequent research expanded to include immigrant student populations in European contexts. More recent studies further broadened the focus to include socio-economically disadvantaged groups, such as students eligible for free school meals in the United Kingdom, as well as specific marginalized populations, such as deaf students in Nepal.

Regarding outcome measures, 17 of the 19 studies assessed academic perfor-

mance. Fourteen studies used course grades or GPA, while eight used standardized test scores. Thirteen studies also measured psychological variables, including school belonging, self-efficacy, perceived identity threat, mental health, and well-being. Six studies collected behavioral data such as attendance rates, disciplinary records, or classroom participation. Follow-up durations ranged from immediate post-tests to long-term assessments spanning over three years.

3.2. Academic Performance Effects

All 19 studies reported data on academic performance, yet the effects varied considerably. 13 studies identified statistically positive effects, four reported no apparent effects, and two observed small negative effects in specific subgroups.

Early studies tended to report stronger outcomes. For instance, Cohen et al. (2006) found that African American students' GPA increased by 0.26 points following the intervention. In a two-year follow-up study, Cohen et al. (2009) observed an increase of 0.24 GPA points among African American students, with gains of 0.18 in the first year and 0.27 in the second year, indicating a cumulative effect over time. Similarly, Sherman et al. (2013) found that Latino students' end-of-year GPA increased by 0.22 points.

However, more recent replication studies produced more conservative findings. Dee (2015) reported nearly effects, while Protzko & Aronson (2016) also found no measurable improvement. By contrast, Borman et al. (2018) conducted a large-scale study and detected a small effect.

Regarding achievement gaps, Cohen et al. (2006) reported that the GPA gap between African American and White students narrowed from 0.75 to 0.45 points. Similarly, Sherman et al. (2013) found that the GPA gap between Latino and White students decreased from 1.02 to 0.80.

Subject-specific analyses revealed that the effects were generally stronger in STEM subjects than in the humanities.

3.3. Psychological Effects

Thirteen studies assessed at least one psychological outcome. Overall, improvements in psychological variables were smaller in magnitude than academic gains but more consistent in terms of statistical significance.

In a large-scale study of 2234 Chinese adolescents, Yan et al. (2023) examined several indicators of psychological well-being. The results showed that life satisfaction improved the most, followed by positive affect, while self-esteem and general mental health showed smaller but still positive changes. Of note, during the COVID-19 pandemic, students in the control group experienced a significant decline in their sense of purpose, whereas those in the affirmation group maintained stability.

School belonging was assessed in eight studies, all reporting improvements following the intervention. In a three-year longitudinal study, Cook et al. (2012) found that students in the affirmation condition maintained a more stable sense of be-

longing compared with the control group, which exhibited a steady decline. Furthermore, the study revealed that low-achieving students in the affirmation condition demonstrated belonging levels comparable to those of high-achieving students in the control condition.

Similarly, Binning et al. (2019) observed that teacher trust declined steadily among students in the control group but decreased at a slower rate among those who received affirmation.

Findings on perceived identity threat were more mixed. Of the five studies that examined this construct, three reported important reductions, while two found no change. Hernandez et al. (2017) demonstrated that self-affirmation alone had limited effect on reducing perceived threat, but when combined with a role-model component, the effect became substantially stronger.

3.4. Behavioral Effects

Six studies reported behavioral outcomes, although the evidence base in this area remains limited.

In terms of disciplinary behavior, Binning et al. (2019) conducted a three-year follow-up and found that annual disciplinary incidents among Black and Latino students in the control group increased from 1.8 to 3.2, whereas the increase in the affirmation group was much smaller, from 1.9 to 2.1. The reduction was particularly notable for more subjective forms of misconduct, such as disrespect or classroom disruption.

Regarding attendance, four studies reported no significant improvement. For example, Dee (2015) found that the average number of absence days was almost identical between the affirmation and control groups (9.2 versus 9.3 days).

Classroom participation was measured in two studies. Cook et al. (2012) reported that students in the affirmation group scored higher on teacher ratings of "answering questions" and "help-seeking" but showed no apparent difference in "homework completion."

3.5. Persistence of Effects

Eleven studies provided follow-up data extending beyond a single academic term, which allowed for an examination of the temporal stability of effects.

Short-term persistence, defined as lasting from one term to one academic year, was consistently observed across all longitudinal studies. Cohen et al. (2009) reported that the effects strengthened progressively over time, with smaller gains in the first term followed by continued increases across subsequent terms.

Evidence of medium-term persistence, defined as lasting one to three years, was identified in five studies. Borman et al. (2018) followed students from Year 7 to Year 9 and found that the positive effects of the intervention persisted throughout the study period. By the end of Year 9, minority students in the affirmation condition had cumulative GPAs that were 0.52 points higher than those in the control group.

Long-term persistence, defined as lasting more than three years, was demonstrated in Goyer et al. (2017), who tracked students from middle school through high school graduation over a period of seven to nine years. The study found several sustained benefits. First, the high-school graduation rate was higher in the affirmation group (94%) than in the control group (87%). Secondly, 89% of Latino students in the affirmation condition enrolled in college-preparatory courses compared with only 50% in the control group. Thirdly, the university entry rate for African American students increased from 78% in the control group to 92% in the affirmation group.

Several mechanisms may explain the persistence of these effects. The accumulation of early success experiences likely enhanced students' confidence, which encouraged greater academic engagement. Increased teacher expectations may also have contributed to more supportive classroom environments. Furthermore, affirmation appeared to foster stronger connections with academically oriented peers, thereby reinforcing positive norms. Over time, these factors may have contributed to a shift in academic self-concept from a sense of inadequacy to a sense of competence.

Nevertheless, attenuation of effects was observed in three studies, primarily when students transitioned to new environments, when very low-performing students lacked the foundational skills needed to benefit from the intervention, or when implementation fidelity was poor.

Finally, in terms of dosage, Cohen et al. (2009) found that adding a second-year reinforcement activity did not produce additional gains, suggesting that a single, well-designed intervention may be sufficient to initiate a self-sustaining process. However, other studies indicated that multiple sessions may yield stronger and more durable outcomes, highlighting the importance of dosage and contextual reinforcement for long-term impact.

4. Discussion

The overall quality of the included studies was moderate, with some variability in findings, suggesting that the conclusions should be interpreted cautiously.

This structured review synthesized evidence from 19 studies examining self-affirmation interventions for adolescents aged 11 - 15 to address five critical research questions: 1) What is the overall effect on educational outcomes? 2) Through what mechanisms do interventions operate? 3) Which factors moderate effectiveness? 4) For whom and under what conditions are benefits most pronounced? and 5) What are the implications for educational practice. The findings reveal that intervention success depends critically on the alignment of multiple psychological, contextual, and implementation factors.

4.1. Overall Effect of Self-Affirmation Interventions on Educational Outcomes

The heterogeneous pattern documented in the results, with two-thirds of studies

showing positive effects and one-third reporting null results, requires interpretation beyond simple vote-counting. Three principal interpretations emerge.

First, heterogeneity most likely reflects genuine contextual dependency rather than methodological noise. Self-affirmation theory predicts effectiveness only when students experience meaningful psychological threat to their academic identity (Cohen & Sherman, 2014; Steele, 1988), not universal benefits across all contexts. In environments where such threat is minimal (because stereotypes are not salient, achievement gaps are narrow, or students possess alternative affirmation sources), benefits are limited. Conversely, where stereotype threat is salient and chronically activated, self-affirmation should provide psychological buffering that enhances outcomes (Cohen et al., 2006, 2009). The observed pattern aligns with this prediction, as studies conducted in contexts with pronounced achievement gaps and salient identity threats more consistently reported positive effects (Cohen et al., 2006, 2009; Sherman et al., 2013; Goyer et al., 2017).

Second, the shift from early strong effects to more mixed recent findings may indicate that initial studies achieved optimal implementation conditions that later studies failed to replicate (Hanselman et al., 2017). Foundational work by Cohen and colleagues demonstrated high implementation fidelity, including subtle delivery, authentic engagement, and precise timing when threat was most salient (Cohen et al., 2006, 2009). As interventions scaled with less researcher involvement, these critical nuances were often lost (Dee, 2015; Borman et al., 2018). Effectiveness thus depends not only on delivering prescribed materials but also on reproducing the psychological conditions that make affirmation authentic and meaningful.

Third, null findings from well-powered studies should be viewed not as failures but as informative evidence delineating the boundaries of effectiveness (Dee, 2015; Bratter et al., 2016; Hanselman et al., 2017). They indicate that self-affirmation is not a universal solution but a context-sensitive intervention effective only under specific conditions (Walton & Yeager, 2020).

It is also notable that effects were more consistent for academic performance and school belonging than for reductions in psychological distress (Yan et al., 2023). This pattern suggests that self-affirmation primarily strengthens adaptive functioning and resource development rather than directly alleviating symptoms of distress, especially in non-clinical populations. The evidence therefore supports positioning self-affirmation as a universal educational support for students facing identity threats rather than as a clinical tool for mental-health improvement.

4.2. Psychological Mechanisms

Clarifying mechanisms is crucial for both theoretical refinement and practical optimization (Cohen & Sherman, 2014). Among proposed mechanisms, the concept of decoupling is most strongly supported (Cook et al., 2012). Self-affirmation disrupts the usual linkage between academic performance and sense of belonging among students vulnerable to stereotype threat. Without affirmation, poor performance predicts diminished belonging, disengagement, and further decline. Af-

firmation interrupts this cycle by enabling students to maintain belonging despite setbacks, preventing temporary academic difficulties from escalating into identity crises (Walton & Cohen, 2007; Cook et al., 2012).

The apparent paradox, namely how belonging improvements yield performance gains, can be explained through recursive processes unfolding over time (Cohen et al., 2009). Sustained belonging fosters engagement, effort, and help-seeking, which improve performance and elicit positive environmental feedback, reinforcing the original psychological change and generating self-amplifying cycles (Cohen et al., 2009; Borman et al., 2019).

Further evidence supports mechanisms of construal broadening and threat reappraisal. Affirmed students demonstrate a greater capacity to interpret challenges abstractly and adaptively (Sherman et al., 2013) and to perceive criticism or difficulty as opportunities for growth rather than as threats to identity (Brady et al., 2016). These processes mitigate stress responses and free cognitive resources for learning (Cohen & Sherman, 2014).

Cross-cultural findings provide an important extension. Successful implementation among Chinese adolescents (Yan et al., 2023) indicates that self-affirmation is not limited to individualistic self-concepts: affirming both individualistic and collectivistic values can restore self-integrity when those values hold personal and cultural meaning (Hoshino-Browne et al., 2005). Nevertheless, many studies still lack rigorous mediation testing, limiting confidence in causal pathways. Future research should employ multiple measurement waves and formal mediation analyses to determine whether observed outcome changes are indeed driven by targeted psychological processes.

4.3. Contextual, Individual, and Implementation Moderators

Context operates as a powerful determinant of effectiveness (Ferrer & Cohen, 2018). At the school level, findings consistently support an optimal threat hypothesis (Hanselman et al., 2014): interventions are most effective when stereotype threat is present but not overwhelming. When minority students are few and gaps are pronounced, affirmation buffers identity threat. When they are the majority, or when discrimination is extreme, interventions may be unnecessary or insufficient (Hadden et al., 2020).

Environmental support further amplifies benefits. The "seed-and-soil" metaphor (Walton & Yeager, 2020) aptly captures this dynamic: psychological seeds of self-integrity flourish only in fertile soil—classrooms with high expectations, supportive teachers, engaged peers, and fair evaluation systems (Borman et al., 2018). In barren contexts, even successful psychological activation cannot translate into sustained behavioral change (Dee, 2015). Peer composition also matters; minority students surrounded by affirming peers often show stronger and more persistent benefits (Binning et al., 2019).

At the individual level, moderate-performing students tend to benefit most, experiencing sufficient threat to make affirmation relevant but not so much that re-

covery feels unattainable (Cohen et al., 2006). Group identification exerts complex influences, since strong identification can intensify threat and thus benefits in hostile settings, but reduce marginal utility in supportive ones (Sherman et al., 2013; Branscombe & Ellemers, 1998).

Implementation quality remains a decisive moderator (Borman et al., 2018). Multi-session formats distributed across time outperform single sessions, which suggests that maintaining self-integrity requires repeated reinforcement (Cohen et al., 2009). Moreover, subtle delivery, by presenting activities as standard classroom tasks rather than remedial programmes, is critical to avoid reactance or stigmatization (Yeager & Walton, 2011). The degree of students' authentic engagement, reflected in emotionally expressive writing, predicts effect magnitude (Borman et al., 2018). Thus, psychological meaning-making, rather than mechanical participation, drives impact.

4.4. Conditions for Maximum Benefit

Synthesizing across these moderators, a coherent contextual model emerges (Walton & Yeager, 2020). The effectiveness of self-affirmation interventions depends on the alignment of four interdependent components that together create the conditions for substantial and sustained impact.

First, meaningful but manageable threat. Students must experience an identity-relevant threat that makes affirmation psychologically salient, yet not so overwhelming that brief interventions cannot provide sufficient buffering (Cohen & Sherman, 2014; Steele & Aronson, 1995). When perceived threat is minimal, affirmation addresses a non-existent problem, and when excessive, psychological resources alone are unlikely to counteract its impact.

Second, environmental affordances. Schools must offer tangible opportunities for affirmed students to demonstrate competence and receive recognition through supportive teachers, rigorous curricula, and fair assessment systems (Cook et al., 2012; Borman et al., 2018). These contextual affordances allow initial psychological changes to evolve into behavioural engagement and academic improvement, ensuring that benefits persist beyond the intervention itself.

Third, quality implementation. Effective delivery requires subtle presentation, appropriate timing, and multiple opportunities for genuine reflection facilitated by well-trained teachers (Cohen et al., 2006; Binning et al., 2019). Interventions achieve their greatest impact when seamlessly integrated into routine classroom activities, enabling students to engage authentically with personally meaningful values rather than perceiving the activity as externally imposed.

Fourth, student receptivity. The students most likely to benefit are those positioned to capitalise on restored self-integrity, namely individuals whose prior performance renders identity threat salient yet recovery attainable (Goyer et al., 2017). Such students possess both the motivational readiness and cognitive capacity for affirmation to trigger recursive improvements in engagement and performance.

When these four components align, self-affirmation can initiate self-reinforc-

ing cycles that strengthen belonging, motivation, and achievement over time (Cohen et al., 2009). Conversely, misalignment in even one component, whether insufficient threat, unsupportive environments, poor implementation, or low student receptivity, can substantially weaken or nullify intervention effects (Dee, 2015; Hanselman et al., 2017).

4.5. Implications for Educational Practice

The evidence yields nuanced implications for practice. Rather than a binary choice between adoption and rejection, schools must engage in strategic, context-sensitive implementation (Cohen & Sherman, 2014).

First, systematic assessment of context should precede adoption. Schools need to determine whether identity-based threats exist and whether environmental conditions can support sustained benefits (Hanselman et al., 2014; Ferrer & Cohen, 2018). Indicators include persistent achievement gaps, student perceptions of stereotyping, or belonging concerns (Walton & Cohen, 2007; Cook et al., 2012). Assessing institutional supports, including teacher expectations, peer cultures, and evaluation fairness, is equally critical, since affirmation cannot substitute for absent structural opportunities (Yeager & Walton, 2011).

Second, psychological interventions must complement rather than replace structural reform. Self-affirmation is most effective when embedded within broader equity initiatives addressing teacher bias, resource inequality, and curriculum disparities (Cohen & Sherman, 2014; Walton & Wilson, 2018). Psychological and structural approaches are mutually reinforcing rather than competing, as the former enables students to navigate challenges while the latter transforms the systems producing them (Borman, 2017).

Third, implementation quality determines outcomes more than protocol fidelity alone (Borman et al., 2018). Effective delivery depends on teacher understanding of underlying psychological principles and on institutional support for continuous monitoring and feedback (Binning et al., 2019). Training should emphasize why subtlety matters, how to foster authentic engagement, and how to maintain cultural and contextual sensitivity (Yeager & Walton, 2011). Continuous quality assurance, via classroom observations, writing analysis, and professional communities, ensures that implementation problems are detected early and addressed promptly (Borman et al., 2018).

Finally, given cross-contextual variability, continuous local evaluation is essential. Schools should monitor achievement, attendance, and belonging before and after implementation, while collecting process data on training, fidelity, and student engagement (Borman et al., 2018). Such monitoring can identify both positive and unintended effects (Dee, 2015), ensuring that interventions remain equitable and effective.

5. Conclusion

Three overarching conclusions can be drawn. Under appropriate conditions, in-

cluding meaningful threat, supportive environments, high-quality implementation, and receptive students, self-affirmation can meaningfully improve educational outcomes (Cohen et al., 2006, 2009; Sherman et al., 2013; Goyer et al., 2017). However, these conditions are not ubiquitous, and effectiveness cannot be assumed from delivery alone (Dee, 2015; Hanselman et al., 2017). Moreover, even under ideal circumstances, self-affirmation addresses only one pathway among many influencing achievement.

The field must therefore move beyond asking whether self-affirmation works toward examining when, how, for whom, and under what conditions it is most effective (Walton & Yeager, 2020). For practitioners, this entails contextually informed decision-making, for researchers, ongoing investigation of moderators, mechanisms, and cultural boundaries, and for policymakers, integration of psychological insights into broader systemic reforms.

Ultimately, self-affirmation represents one promising, low-cost tool within a comprehensive equity framework. Judged by realistic standards, namely producing measurable benefits under specified conditions at reasonable cost, it merits continued research, cautious application, and sustained integration with structural strategies that together advance educational equity.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

Ahn, E., & Kang, H. (2018). Introduction to Systematic Review and Meta-Analysis. *Korean Journal of Anesthesiology, 71,* 103-112. https://doi.org/10.4097/kjae.2018.71.2.103

Binning, K. R., Cook, J. E., Purdie-Greenaway, V., Garcia, J., Chen, S., Apfel, N. et al. (2019). Bolstering Trust and Reducing Discipline Incidents at a Diverse Middle School: How Self-Affirmation Affects Behavioral Conduct during the Transition to Adolescence. *Journal of School Psychology, 75,* 74-88. https://doi.org/10.1016/j.jsp.2019.07.007

Borman, G. D. (2017). Advancing Values Affirmation as a Scalable Strategy for Mitigating Identity Threats and Narrowing National Achievement Gaps. *Proceedings of the National Academy of Sciences, 114,* 7486-7488. https://doi.org/10.1073/pnas.1708813114

Borman, G. D., Grigg, J., Rozek, C. S., Hanselman, P., & Dewey, N. A. (2018). Self-Affirmation Effects Are Produced by School Context, Student Engagement with the Intervention, and Time: Lessons from a District-Wide Implementation. *Psychological Science, 29,* 1773-1784. https://doi.org/10.1177/0956797618784016

Borman, G. D., Rozek, C. S., Pyne, J., & Hanselman, P. (2019). Reappraising Academic and Social Adversity Improves Middle School Students' Academic Achievement, Behavior, and Well-Being. *Proceedings of the National Academy of Sciences, 116,* 16286-16291. https://doi.org/10.1073/pnas.1820317116

Brady, S. T., Reeves, S. L., Garcia, J., Purdie-Vaughns, V., Cook, J. E., Taborsky-Barba, S. et al. (2016). The Psychology of the Affirmed Learner: Spontaneous Self-Affirmation in the Face of Stress. *Journal of Educational Psychology, 108,* 353-373. https://doi.org/10.1037/edu0000091

Branscombe, N. R., & Ellemers, N. (1998). Coping with Group-Based Discrimination: Individualistic versus Group-Level Strategies. In J. K. Swim, & C. Stangor (Eds.) *Prejudice*

(pp. 243-266). Elsevier. https://doi.org/10.1016/b978-012679130-3/50046-6

Bratter, J. L., Rowley, K. J., & Chukhray, I. (2016). Does a Self-Affirmation Intervention Reduce Stereotype Threat in Black and Hispanic High Schools? *Race and Social Problems, 8,* 340-356. https://doi.org/10.1007/s12552-016-9187-4

Cohen, G. L., & Sherman, D. K. (2014). The Psychology of Change: Self-Affirmation and Social Psychological Intervention. *Annual Review of Psychology, 65,* 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,* 1307-1310. https://doi.org/10.1126/science.1128317

Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., & Brzustoski, P. (2009). Recursive Processes in Self-Affirmation: Intervening to Close the Minority Achievement Gap. *Science, 324,* 400-403. https://doi.org/10.1126/science.1170769

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,* 479-496. https://doi.org/10.1037/a0026312

Covarrubias, R., Herrmann, S. D., & Fryberg, S. A. (2016). Affirming the Interdependent Self: Implications for Latino Student Performance. *Basic and Applied Social Psychology, 38,* 47-57. https://doi.org/10.1080/01973533.2015.1129609

Creswell, J. D., Dutcher, J. M., Klein, W. M. P., Harris, P. R., & Levine, J. M. (2013). Self-Affirmation Improves Problem-Solving under Stress. *PLOS ONE, 8,* e62593. https://doi.org/10.1371/journal.pone.0062593

Dee, T. S. (2015). Social Identity and Achievement Gaps: Evidence from an Affirmation Intervention. *Journal of Research on Educational Effectiveness, 8,* 149-168. https://doi.org/10.1080/19345747.2014.906009

Escobar-Soler, C., Berrios, R., Peñaloza-Díaz, G., Melis-Rivera, C., Caqueo-Urízar, A., Ponce-Correa, F. et al. (2024). Effectiveness of Self-Affirmation Interventions in Educational Settings: A Meta-Analysis. *Healthcare, 12,* Article 3. https://doi.org/10.3390/healthcare12010003

Ferrer, E., & Cohen, G. L. (2018). A Putative Role for Lay Theories in the Effectiveness of Psychological Interventions. In C. Dweck, D. S. Yeager, & G. M. Walton (Eds.), *Handbook of Competence and Motivation* (2nd ed., pp. 620-643). Guilford Press.

Gorard, S. (2024). Judging the Relative Trustworthiness of Research Results: How to Do It and Why It Matters. *Review of Education, 12,* e3448. https://doi.org/10.1002/rev3.3448

Goyer, J. P., Garcia, J., Purdie-Vaughns, V., Binning, K. R., Cook, J. E., Reeves, S. L. et al. (2017). Self-Affirmation Facilitates Minority Middle Schoolers' Progress along College Trajectories. *Proceedings of the National Academy of Sciences, 114,* 7594-7599. https://doi.org/10.1073/pnas.1617923114

Hadden, I. R., Easterbrook, M. J., Nieuwenhuis, M., Fox, K. J., & Dolan, P. (2020). Self-Affirmation Reduces the Socioeconomic Attainment Gap in Schools in England. *British Journal of Educational Psychology, 90,* 517-536. https://doi.org/10.1111/bjep.12291

Hanselman, P., Bruch, S. K., Gamoran, A., & Borman, G. D. (2014). Threat in Context: School Moderation of the Impact of Social Identity Threat on Racial/Ethnic Achievement Gaps. *Sociology of Education, 87,* 106-124. https://doi.org/10.1177/0038040714525970

Hanselman, P., Rozek, C. S., Grigg, J., & Borman, G. D. (2017). New Evidence on Self-Affirmation Effects and Theorized Sources of Heterogeneity from Large-Scale Replica-

tions. *Journal of Educational Psychology, 109,* 405-424. https://doi.org/10.1037/edu0000141

Hernandez, P. R., Bodin, R., Haeger, H., Cabay, M., Graham, K., & Coe, K. (2017). Tightening the Gap: A Self-Affirmation Intervention to Boost the Academic Performance of Latino Students in a High-Stakes Course. *Journal of Educational Psychology, 109,* 1059-1073.

Hoshino-Browne, E., Zanna, A. S., Spencer, S. J., Zanna, M. P., Kitayama, S., & Lackenbauer, S. (2005). On the Cultural Guises of Cognitive Dissonance: The Case of Easterners and Westerners. *Journal of Personality and Social Psychology, 89,* 294-310. https://doi.org/10.1037/0022-3514.89.3.294

Lokhande, M., & Müller, T. (2019). Double Jeopardy—Double Remedy? The Effectiveness of Self-Affirmation for Improving Doubly Disadvantaged Students' Mathematical Performance. *Journal of School Psychology, 75,* 58-73. https://doi.org/10.1016/j.jsp.2019.07.006

Protzko, J., & Aronson, J. (2016). Context Moderates Affirmation Effects on the Ethnic Achievement Gap. *Social Psychological and Personality Science, 7,* 500-507. https://doi.org/10.1177/1948550616646426

See, B. H., Morris, R., Gorard, S., Siddiqui, N., Easterbrook, M. J., Nieuwenhuis, M. et al. (2022). A Conceptual Replication Study of a Self-Affirmation Intervention to Improve the Academic Achievement of Low-Income Pupils in England. *Educational Research and Evaluation, 27,* 83-116. https://doi.org/10.1080/13803611.2021.2022317

Sherman, D. K., Hartson, K. A., Binning, K. R., Purdie-Vaughns, V., Garcia, J., Taborsky-Barba, S. et al. (2013). Deflecting the Trajectory and Changing the Narrative: How Self-Affirmation Affects Academic Performance and Motivation under Identity Threat. *Journal of Personality and Social Psychology, 104,* 591-618. https://doi.org/10.1037/a0031495

Steele, C. M. (1988). The Psychology of Self-Affirmation: Sustaining the Integrity of the Self. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 21, pp. 261-302). Elsevier. https://doi.org/10.1016/s0065-2601(08)60229-4

Steele, C. M., & Aronson, J. (1995). Stereotype Threat and the Intellectual Test Performance of African Americans. *Journal of Personality and Social Psychology, 69,* 797-811. https://doi.org/10.1037/0022-3514.69.5.797

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., & Wilson, T. D. (2018). Wise Interventions: Psychological Remedies for Social and Personal Problems. *Psychological Review, 125,* 617-655. https://doi.org/10.1037/rev0000115

Walton, G. M., & Yeager, D. S. (2020). Seed and Soil: Psychological Affordances in Contexts Help to Explain Where Wise Interventions Succeed or Fail. *Current Directions in Psychological Science, 29,* 219-226. https://doi.org/10.1177/0963721420904453

Wu, Z. (2024). *Understanding the Effects of Values Affirmation on the Academic Performance and Psychological Well-Being of Deaf Students in Nepal: Evidence from a Randomized Controlled Trial.* Ph.D. Thesis, New York University.

Yan, W., Wang, Y., Jiang, Z., Peng, K., & Cohen, G. (2023). Effects of a Self-Affirmation Intervention among Chinese Adolescents during the COVID-19 Pandemic: A Large-Scale Randomized Controlled Trial. *Applied Psychology: Health and Well-Being, 16,* 1100-1121. https://doi.org/10.1111/aphw.12516

Yang, J., Zhao, Y., Ji, Y., Ma, J., Li, L., & Hu, X. (2023). Randomised Controlled Trial of Self-Affirmation Intervention on Students' Academic Performance: Promising Impacts

on Students from Migrant Hukou Status. *Psychology Research and Behavior Management, 16,* 3607-3621. https://doi.org/10.2147/prbm.s419112

Yeager, D. S., & Walton, G. M. (2011). Social-Psychological Interventions in Education: They're Not Magic. *Review of Educational Research, 81,* 267-301. https://doi.org/10.3102/0034654311405999