Who benefits from the psychology of achievement? The effectiveness and equity of school-based psychosocial interventions for promoting academic resilience to improve educational performance

Stephanie Osahenrunmwen Odiase

MPhil Candidate in Evidence-Based Social Intervention and Policy Evaluation
Department of Social Policy and Intervention
University of Oxford

Date: 29.06.2018

Word Count (Text): 27,452
Word Count (Appendices, Tables, Figures): 5,254
Word Count (Abstract): 481

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

0 Abstract

Background. Academic underperformance is a serious concern affecting students worldwide. Educational inequity contributes to academic underperformance whereby students with specific sociodemographic characteristics receive less school resources and adequate teaching relative to their peers. Gender, race, and socioeconomic status act as salient and common sources of educational inequity, and as such, serve as risk factors that increase the likelihood of academic failure. Academic resilience, defined as the increased likelihood of school success in the face of adversities, has been proposed as a potential solution for the deleterious effects of academic risk. Fostering academic protective factors through school-based interventions to improve academic resilience may promote effective and equitable school performance.

Objectives. The study had two main objectives. Firstly, the review aimed to determine if gains in academic resilience via the use of specific psychosocial protective factors can improve academic performance among primary and secondary school students. The second objective was to assess whether or not relative improvements in school performance were equitable across the PROGRESS-Plus factors of gender, race, and socioeconomic status.

Methods. A comprehensive search for prominent academic-related psychosocial protective factors informed the search strategy for relevant studies. A systematic review of school-based interventions for academic resilience was conducted, including database searching, grey literature searching, and contacting relevant authors for additional research. Effectiveness data syntheses were performed using a meta-analysis of main effects across ten studies as well as a sensitivity analysis for two low-risk of bias studies. Harvest plots synthesised equity evidence across four studies while equity-relevant information was

summarised in a bar graph. Several robustness checks (funnel plot, Egger's test, influence diagnostics) were performed as an explanatory measure to account for sources of bias and between-study variability.

Results. The results indicated that academic resilience interventions are effective at ameliorating school performance (g=0.8, p=0.008). The equity results from the harvest plot suggest that these interventions may reduce educational inequity across gender and race for female students and minority students but no evidence was reported on socioeconomic status. The results from the sensitivity analysis were large (g=2.11) but statistically insignificant (p=0.19). The robustness checks confirmed the presence of statistical heterogeneity, which was attributed to the inclusion of one study, Miller (2017).

Conclusions. The findings indicated that academic resilience interventions may be both effective and equitable at improving educational performance. Given the novelty of this research field, additional research is needed across all academic protective factors, especially self-care behaviours and academic help-seeking. Additionally, better reporting standards, including equity guidelines for non-randomised trials and published protocols, are vital to future research endeavours. Academic resilience interventions have major practice implications given the flexibility in delivery, wide range of use across differing school levels, and cross-cultural relevance. Considering international education policy, these interventions provide an insightful and innovative way of closing global educational gaps, which may have large implications for the U.N. Sustainable Development Goal 4 for providing high-quality and accessible education for all.

1 Background

1.1 Origins of academic underperformance and inequality

Academic underperformance is a serious global concern impacting the contemporary international landscape. Thirty-eight percent of children worldwide who have completed their primary education have not received the most basic learning skills needed to succeed in life (UNESCO, 2014). In the United States, students are trailing far behind in the international mathematics rankings, even when adjusting for the greater educational spending per child in comparison to most countries (OECD, 2012). Of tantamount, if not greater, importance is the issue of educational inequality, as disparities in education may lead to underperformance in school for certain children. The unequal distribution of academic resources (i.e. trained teaching staff, funding, textbooks, school lunches, etc.) across communities is a *contributor* to educational inequality. However, this can be seen as a *consequence* of educational inequality as well, stemming from the larger and more pervasive issues of racial, socioeconomic, and gender differences that typically result in historically disadvantaged children receiving less than their peers (Peske & Haycock, 2006; The Centre for Social Justice, 2014).

Extensive literature has demonstrated that different facets of a young person's background, such as his or her gender, socioeconomic status, race, and many other factors, can create potentially undesirable social gradients in educational attainment (Bradley, 2014; Ladson-Billings & Tate, 1995; Lynch & O'Riordan, 1998; Stromquist, 1990; Troyna, 2012). For example, considering the implications of between-country socioeconomic differences, two percent of adolescents do not attend school in high income countries versus the staggering 36% of adolescents living in low income countries (UNESCO, n.d.). Even in a

lower-middle income country such as Nigeria, educational attainment varies greatly within region by wealth. Three quarters of the poorest children aged 7 to 16 years in the Northeast region have never attended school, whereas virtually all of the richest children within the same age group and location have, demonstrating the within-country effects of these social determinants of education (UNESCO, 2013b; World Bank, 2016).

1.2 Description of the condition

1.2.1 Social determinants of education

The social determinants of education present a complex and interrelated conceptual framework explaining for whom and under which circumstances some students face academic adversity. Borrowing from the World Health Organisation's Commission on the Social Determinants of Health, their model has been extended beyond its epidemiological origins and applied in an educational context in order to inform the theoretical underpinnings of this equity review (Marmot & Health, 2008; Tugwell et al., 2010). Though there are various causes for academic disparity, including disability status and language barriers (Parker, Rubalcava, & Teruel, 2005; Powell, 2003), race, socioeconomic status, and gender are all of particular significance as these three factors have the largest impact on academic performance.

1.2.2 Socioeconomic inequality

The term 'socioeconomic status' (SES) refers to the combination of economic and sociological elements, like household income or parents' level of education, which as a whole, serves as a stratifying factor in society, and matters insofar as it is a plausible explanation for academic shortcomings among certain students. Socioeconomic status is intrinsically linked to educational inequality, primarily because it is a strong predictor of a

child's place of residence. In general, this can have a negative impact on catchment areas, and thus the type of school children will attend if they are located in lower-income settings. If a school is in a disadvantaged area, then it is more likely to lack the proper funding and resources needed to effectively teach students (Aikens & Barbarin, 2008). American and British students are somewhat less restricted by catchment areas. This is due to national policy structures like the No Child Left Behind Act of 2001 and the British government allowing families to apply to schools in other local councils. However, school allotment by catchment area is typically the case in most countries, and as such, may be a compromising factor for most students.

Additionally, children raised in low-income settings have parents who themselves were likely raised in similar conditions given the cyclic nature of poverty (Valentine, 1968). Poverty trap theory posits that families are restricted by intergenerational social and economic drawbacks, so social mobility and access to better resources are hindered, thus limiting educational opportunities (Barham, Boadway, Marchand, & Pestieau, 1995; Bowles & Gintis, 2002; Crosnoe, Mistry, & Elder, 2002). Consequently, it is harder for low SES students to break the vicious circle of poverty and attain a worthwhile education. This results in students who become academically handicapped, in school as well as at home, as children's scholastic abilities are highly correlated with their home literacy environment and parent's level of education (Davis-Kean, 2005; H. B. Ferguson, Bovaird, & Mueller, 2007; Hirsch, 2007; van Bergen, van Zuijen, Bishop, & de Jong, 2017).

1.2.3 Racial inequality

Race is considered to be a social construct categorising people into discrete groups, such as White, Black, Latino, mixed race, etc. Although ethnicity differs slightly from race in terms of encompassing certain socio-cultural features, such as language, nationality, and

regional culture (Fishman, 1997), due to the overall similarity of these concepts, they will be used interchangeably throughout this review. The impact of educational inequality is worsened by the strong relationship between race and class, as low levels of academic achievement among poor ethnic minority students is an often-explored phenomenon across sociology, psychology, and education (Rigsby, Stull, & Nancy, 1997). Considering that minority children have a higher probability of growing up in concentrated and persistent poverty, they are more likely to attend resource-poor schools, and thus experience academic failure (Brooks-Gunn, Duncan, & Aber, 1997). Research suggests that school racial composition and the presence of qualified teaching staff are strongly associated with performance scores. Students in schools with higher percentages of minorities and economic instability do not have as many trained instructors, and thus do not fare well academically relative to their Caucasian peers (Orfield & Lee, 2005). As a consequence of the interplay between race, socioeconomic status, and academic achievement, failure to complete schooling is more common among low-income minority students, with a projected upward trend of 50% of poor African American and Latino students dropping out of high school in the United States (Orfield, 2004).

The psychology of race can play a role in educational inequality as well. Theories on stereotype threat suggest that the *perceived* abilities of minority students, such as Black, Latino, and Native American, are lower than those of White and Asian students (Steele, 1992, 1997). The processes of stereotype threat can unfold in a number of ways. Given the implicit bias against certain races by teachers and institutions, the imposition of negative stereotypes about ethnic students' intellectual capacities may induce higher levels of stress, especially among those who identify more with this stereotype. Furthermore, minority students may also conform to negative stereotypes through stigma consciousness, thus aligning one's self-concept with academic failure (Brown & Pinel, 2003; Guyll, Madon,

Prieto, & Scherr, 2010; Steele, 1997). These causes may impact immediate academic performance, such as test-taking, as well as beget chronic anxiety, the culmination of which may result in both academic de-motivation and withdrawal (i.e. dropping out) (Osborne & Walker, 2006; Steele, 1997; Steele & Aronson, 1995). This is further complicated by the negative socialisation that ethnic students experience at the school level via teachers and non-ethnic peers who underrate their performance (Jacob, 1981; Orfield & Lee, 2005). Combined with race and racial discrimination theory in educational sociology, this may explain the higher educational system levels of discrimination in terms of allocation of school resources among ethnic students, especially those of African/Caribbean origin (Coard, 1971; Troyna, 1984, 1991). Thus, racial discrimination above and beyond socioeconomic gradients may produce differential effects in academic performance as some evidence from the United States suggests that even lower class White students may perform better than higher SES Black students (Miller, 1997).

1.2.4 Gender inequality

Gender, within the scope of this review, classifies students by preference as opposed to biological assignment as either 'female' or 'male,' and has been as defined as such, with the assumption that a negligible percentage of students will self-identify as 'transgender' in the included studies. The relationship between gender inequality and academic achievement is a deeply complex matter. How gender inequality manifests in education varies greatly by context, as it can interplay with geographical location, race, academic subject, and global trends. As such, understanding the underpinnings of gender that beget educational inequality can be conceptually intricate. Similar to race, research has asserted that educational inequalities can arise due to stereotype threat by gender (Beilock, 2008). Especially with respect to performance in mathematics, gender stereotypes reinforce

the myth that female students lack the mental acuity needed to do as well as their male peers. For example, research from the United States suggests that teachers' perceptions can moderate female students' capabilities such that negative beliefs may prompt poor academic performance (Beilock, Gunderson, Ramirez, & Levine, 2010). It is important to highlight the subject-specific nature of these findings as female students perform as well if not better than their male peers in science across 30 countries (National Center for Educational Statistics, n.d.). Thus, scholastic ability is impeded not by innate talent, or lack thereof, but rather by the threat that their performance may confirm the negative gender stereotype (Tomasetto, Alparone, & Cadinu, 2011).

Of further concern is the age at which stereotype threat begins to erode scholastic achievement. Evidence indicates that female students as early as the first grade exhibit an aversion to mathematics. This is perhaps in response to the burgeoning awareness of their gender identity, and thus the potentially unknown gender-normative behaviours that may follow suit (Ambady, Shih, Kim, & Pittinsky, 2001). This theoretical framework of inequality may validate the Programme for International Student Assessment (PISA) results for secondary school students in mathematics in which girls outperform boys in only five out of 68 countries, whereas male students lead in 41 of 68 countries (Cook, 2014; OECD, 2012).

The plight of gender inequality in education is far worse in developing countries, as girls are far less likely than boys to have basic access to schools (UNESCO, 2013a). More sources of academic disparity arise when young women face additional challenges such as pregnancy, which is a particular barrier to education in this context. Childbirth can affect the education of young women through multiple pathways. There are macro-level consequences for public resources due to increasing populations, which cause strain on

local school systems in terms of resources, funding, and space (Jenkins, 2015). Additionally, household-level resources are diluted as raising more children may divert vital funding from schooling costs (Blake, 1989). The most immediate, and perhaps most damaging, outcome is the increased likelihood of teenage mothers dropping out of school, with pregnancies accounting for 18% of female dropouts in secondary schools among 23 sub-Saharan countries (DHS, 2003; Klepinger, Lundberg, & Plotnick, 1995). What results is that fewer women obtain an education, limiting their employment opportunities and thus spending more time childrearing. This chain of events may compromise the next generation as well, as better-educated mothers are more likely to support their daughter's education, and those without support and a robust home-learning environment are more likely to perpetuate these academic shortcomings (Lloyd & Blanc, 1996; UNFPA, 2002).

However, the impediment of gender switches from female to male when accounting the effects of race in high-income countries like the United States. Building on the nature of stereotype threat conveying misguided beliefs that ethnic male students, especially those of Black descent, lack intellectual competence is the major threat of these students being branded as socially belligerent (Hudley & Graham, 2001; Steele, 1997). This additional burden creates a false notion that African-American males are not only 'unintelligent,' but that they also are hard to teach given their assumed maladjusted behaviour. What emerges from this structurally-imposed racism is oppositional behaviour, stemming from Ogbu's theory on oppositional culture (Ogbu, 1991). Given the experience of racial disadvantage, Black male students may possess high levels of cultural mistrust and therefore expect institutions such as schools to treat them in an unfair manner (Ogbu, 1991). This mistrust is harmful to academic progress, as these students begin to exhibit high levels of deviant behaviour given their dissatisfaction with their educational milieu (Biafora et al., 1993; Irving & Hudley, 2005). Furthermore, Black male students may have

lowered their expectation of success, and as a result, attribute less value to the ideals of academic accomplishment because they may believe that they will not reap any personal benefits from receiving an education (Irving & Hudley, 2005).

Borrowing from cultural ecological theory, the culmination of cultural mistrust, oppositional behaviour, expectations of educational outcomes, and perceived academic value have sequestered some Black male students. These students may believe that they cannot access learning opportunities, and even if they could, it would be frowned upon because the value of obtaining an education is culturally aligned with the forces that may have ostracised them in the first place, such as educational systems and mainstream society (Irving & Hudley, 2005; Ogbu & Simons, 1998). As a consequence, across all levels of education, Black male students are far behind both Black female and White male students in most societies (Ferguson, 2003; Hrabowski III, Maton, & Greif, 1998; Polite & Davis, 1999). Black male students are also the most likely group to be expelled or suspended from school (Meier, Stewart, & England, 1998), underrepresented in gifted or advanced placement programmes (Grantham, 2004), overrepresented in remedial or special education classes (Milofsky, 1974), and disengaged academically (Ford, 2011).

The aforementioned conceptual frameworks on the social determinants of education are not meant to be an exhaustive depiction. Rather, these noted characteristics and their related theories underscore the complexity and variability of the many causes of educational inequality. Furthermore, elucidating these theories pinpoint several specific risk factors that undermine educational achievement, thus highlighting circumstances in which school-based interventions can improve academic risk.

1.3 How the concepts of risk and resilience inform school-based research

1.3.1 Conceptualising academic risk

The *adverse experiences* that stem from these social determinants of education, and not the *features* themselves (i.e. race, gender, etc.), act as risk factors that can hamper a child's development, and thus their academic progression. Risk typically refers to 'conditions that could pose a significant threat to the adaptive function or development of a person' (Masten, 2013, p. 2). Related to the concept of risk are risk factors, which are measurable and observable characteristics which may predict negative outcomes (Masten, 2013). There are various types of risk factors, including biological indicators (e.g. congenital brain defects), sociodemographic factors (e.g. socioeconomic status, race), and life experiences (e.g. sexual abuse, loss of a loved one), all of which are highly associated with maladaptive outcomes (Masten, Best, & Garmezy, 1990). It is important to stress that risk is a probabilistic concept and that exposure to certain adversities does not guarantee that a child will become vulnerable since the same risk factor can affect children differently (Belsky, Bakermans-Kranenburg, & Van Ijzendoorn, 2007).

Figure 1. Bronfenbrenner's (1979) ecological systems model developed with context-specific risk factors that may impede academic performance. A black circle has to added to the model to signify that the interventions included in this review target the individual-level and micro-level risk factors (excluding family) given the focus on school-based interventions.

Risk factors can impede a child's academic competence at various levels of his or her social surroundings. As illustrated in figure 1 above, ecological systems theory outlines the environmental systems with which children interact, and it is within these contexts that the risk factors that compromise academic achievement may arise (Bronfenbrenner, 1979). Though the included studies for this review target academic protective factors at the individual and school levels, all levels of the Bronfenbrenner model are showcased to fully

illustrate the various interrelated risks that can negatively impact a child's scholastic achievement. For example, individual-level risk factors could include the aforementioned social determinants of education like gender and race, and microsystem-level settings such as a poor home environment could increase academic risk due to lack of literacy tools (e.g. children's books) and low-income parents who are less likely to mentally engage with their children (Farver, Xu, Eppe, & Lonigan, 2006; Magnuson, Sexton, Davis-Kean, & Huston, 2009). Coupled with resource-poor schools in low-income areas, the mesosystem-level interactions between the home and formal educational institutions can also produce risk factors due to low SES status. Exosystem-level risk can emerge through a dearth of social services, which has a trickle-down effect, as these services may be unavailable to students due to the lack of local funding, thus compromising a child's educational opportunities further. The macrosystem encompasses sociocultural disadvantages such as abject poverty and racism that systematically disenfranchise certain members of society, like impoverished students in developing countries (UNESCO, 2013b).

When these risk factors combine, they increase a child's cumulative risk indices. This produces a risk gradient wherein the likelihood and number of maladaptive outcomes are directly related to the number of psychosocial risk factors, leading to the possibility for these vulnerabilities to accumulate in such a way that has a compounded negative effect (Appleyard, Egeland, van Dulmen, & Sroufe, 2005; see figure 2 below). Connecting this logic with a conceptual framework, developmental cascades theory suggests that the cumulative consequences across the levels alter the course of development which has a specific impact on general cognitive competence, often measured by academic achievement or intellectual ability (Kohlberg, LaCrosse, & Ricks, 1972). Thus, a general causal pathway of academic risk would assume that risks factors, and especially their cumulative impact, could weaken students' psychosocial skills, without which they do not possess the

necessary adaptive behaviour to persevere through school, leading to academic underperformance. For example, being a poor, female, ethnic minority student in a lower-income country would create many barriers to receiving an education. Given the number of adversities, this student is more likely to lack the psychosocial resources needed to overcome these barriers in comparison to a student with fewer risk factors, thus dramatically hindering her educational success.

Accounting for risk, especially psychosocial risk factors, in academic achievement is essential as it is a measure of resources, because the presence of an risk factor could indicate the lack of a vital asset (i.e. protective factor), like a lack of social support could prompt academic failure (Leary & DeRosier, 2012; Masten, 2013). Additionally, acknowledging and assessing risk may lead to a better understanding of its causes. Ethnicity, for example, could be a proxy measure for the actual risk factor of low socioeconomic status, which could lead to identifying the additional threats of poor schooling and inadequate healthcare (e.g. not having glasses may lead to a child being wrongly placed in remedial classes) (Masten, 2013). Especially when considering the gravity of cumulative risk indices and their ability to precipitate negative developmental cascades, it is important to better understand when and how the exact processes in a child's development give rise to worsened outcomes. This includes pinpointing when the risks that negatively impact academic competence will start to have lasting effects on the future (Masten, 2013).

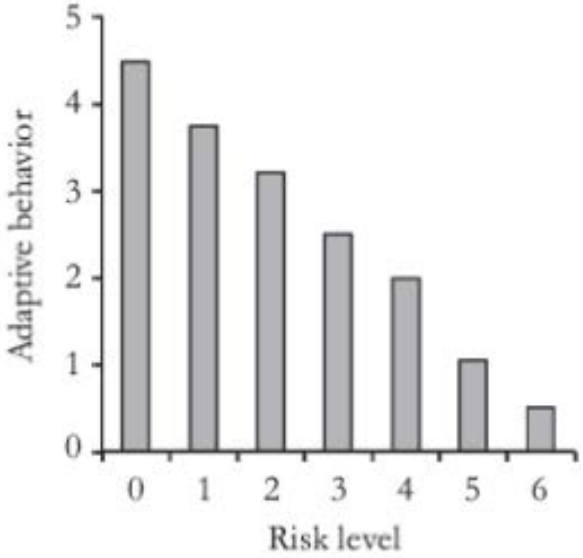


Figure 2. Masten's (2013) gradient of cumulative risk indices

1.3.2 Conceptualising academic resilience

Resilience is 'an elusive construct' (Borman & Overman, 2004, p. 193). This is due in part to the lack of consensus on the domains of the construct, as well as a standardised typology of adaptive behaviours that resilience encompasses (Masten, 1994). This may explain why there are numerous definitions for resilience throughout the literature (Gordon, 1995; Luthar, Cicchetti, & Becker, 2000; Masten et al., 1990; Rutter, Maughan, Mortimore, Ouston, & Smith, 1979; Werner & Smith, 1992). Given the level of inconsistency on what resilience is and how to operationalise it in research, this concept remains difficult to effectively measure and assess in social intervention programmes, undercutting its scientific value. Thus, researchers wonder whether it is more beneficial to have a context-specific approach to evaluating resilience in lieu of a global, unidimensional construct which can be conceptually vague (Liddle, 1994; Riley & Masten, 2005). As such, academic resilience (also commonly referred to *academic resiliency* or *educational resilience*) serves a context-specific version of resilience that affords greater assessment and predictive validity in this field of research (Cassidy, 2016; Colp & Nordstokke, 2014).

Academic resilience has been defined as the increased likelihood of school success and other related life accomplishments, in the face of adversities in one's environment, which stems from early traits, conditions, and experiences (M. C. Wang, Haertel, & Walberg, 1994). However, the review author contends that academic resilience does not only stem from early life characteristics but can also be developed throughout the academic life course through the use of school-based interventions, as this review aims to test. Academic resilience is a sounder theoretical construct than global resilience and as such, it is a more robust means of pinpointing the psychosocial tools needed to assist vulnerable students at risk of failure (Cassidy, 2016).

1.3.3 Defining academic protective factors

If academic resilience is the mechanism needed to reach academic success among at-risk students, then protective factors are the tools behind the mechanism. Analogous to the relationship between risk and risk factors, protective factors are measurable and observable characteristics that increase the probability of a student becoming resilient and having positive outcomes despite their vulnerabilities (Masten, 2013). For the purpose of school-based interventions, protective factors can be viewed as the key programme components, or 'active ingredients,' that lead to academic resilience. An essential, and sometimes overlooked, point of clarification is that protective factors (also known as 'assets') matter insofar as there is risk present. Simply put, students can only be academically resilient if and only if they face chronic and persistent risk as members of a 'disadvantaged' group (Morales & Trotman, 2004). This is not to say that other students who are not considered 'at-risk,' such as healthy middle-class Caucasian children, cannot have instances where they experience educational shortcomings. However, in this case, the uptake of adaptive behaviours to offset academic underperformance would be classified as 'promotive factors' which would foster *academic buoyancy* (Martin & Marsh, 2009; Masten, 2013). These concepts and relevant interventions are different since the peaks and valleys every student may experience are less severe than the ongoing adversities that would necessitate academic resilience (Martin & Marsh, 2009).

Given the interest in targeting and promoting psychosocial protective factors that improve educational achievement, it was necessary to better understand which specific attributes foster academic resilience, and thus their resulting impact on school performance. A literature search was conducted to identify as many academic-related psychosocial protective factors as possible for the purposes of informing the theoretical

basis of this review. This was done by searching electronic databases (ERIC, Google Scholar, etc.). Relevant studies were then reference-checked for additional literature. More information was gathered from the book *Promoting Academic Resilience in Multicultural America* (Morales & Trotman, 2004) and contacting relevant researchers in the field of academic resilience (see appendix 7.7). The search for relevant psychosocial protective factors concluded once data saturation was reached, i.e. themes began to repeat in the literature and correspondence.

The identified psychosocial protective factors were the following: Academic motivation, self-efficacy and self-regulation, coping, school-based social support, academic perseverance, academic planning, adaptive help-seeking, self-care behaviours, and cognitive style.

Academic motivation is defined as the desire to perform well in school, which is influenced by a student's perceived value in his or her educational pursuits, as well as the expectation that maintaining a high level of engagement will produced positive results (Martin & Marsh, 2006; Pajares, 2001). If academic motivation can be conceptualised as the desire to do well in school, then academic planning can be seen as the actions that stem from said desire. It specifically relates to a student's ability to coordinate relevant behaviour with his or her goal orientation (i.e. goal setting) to accomplish a predetermined task (e.g. improving one's grades in a specific course) (Jowkar, Kojuri, Kohoulat, & Hayat, 2014; Martin & Marsh, 2006). It should be noted that academic planning and self-regulation are similar constructs, but self-regulation, defined below, refers more to engaging in regular positive educational practices whereas academic planning is typically done with a specific goal in mind.

Academic self-efficacy refers to a student's judgements of his or her own capabilities to succeed academically (Bandura, 1997). It should be noted that it is not the

same as self-worth, which is more akin to self-concept (see 'cognitive style'), but rather the degree of confidence to which students believe that they possess the necessary skills to accomplish their educational goals (Martin & Marsh, 2006; Pajares, 2001). Self-regulation, a related theoretical construct, is the behaviour behind the beliefs of self-efficacy. Self-regulatory strategies include finishing schoolwork within an appropriate amount of time, regularly practicing materials outside of one's academic comfort zone, etc. (Nota, Soresi, & Zimmerman, 2004; Zimmerman & Schunk, 2001). Self-regulation, in regards to self-regulated learning, also encompasses the self-reflective practice of understanding which skills are needed to improve scholastic achievement (i.e. thinking about learning effectively) and the necessary steps to personally strengthen any metacognitive academic weaknesses (Hu et al., 2011; Schunk & Zimmerman, 1998).

Coping is the combination of personal, social, and behavioural abilities that affect the manner in which students comfort themselves stressful academic situations (Leary & DeRosier, 2012). Self-care behaviours represent the extent to which students engage in health-promoting behaviours during times of stress and the ways in which these actions positively influence psychoeducational functioning (Leary & DeRosier, 2012). School-based social support reflects the level of meaningful involvement and interaction that a student has with his or her school environment, which leads to better integration, and thus increased educational performance (Morales & Trotman, 2004). School-based social support can be instilled in a number of ways, including words of encouragement, peer connectedness, support from teachers and other school authorities, mentorship groups, and having a nurturing school environment (Johnson, Taasoobshirazi, Kestler, & Cordova, 2015; Morales, 2000).

Academic perseverance represents the willingness and sheer determination to persist through academically challenging experiences that may expose students to

prolonged periods of chronic stress (Martin & Marsh, 2006). This asset extends above and beyond the simple fervour of wanting to perform well in school (i.e. academic motivation) and as such, it most closely resembles general resilience. Adaptive help-seeking refers to the process by which students to identify their own learning obstacles and then actively seek out assistance from others, thereby ameliorating their academic performance (Cassidy, 2016).

Finally, cognitive style embodies a variety of thoughts, beliefs, and attitudes through which students attribute value to their own abilities and their outlook on their academic future. Especially significant to this protective factor is the notion of self-concept, which includes internal locus of control (Leary & DeRosier, 2012). Students who possess internal loci of control intrinsically believe that their academic successes and failures are within their grasp, and that they have the ability to direct their educational experiences and outcomes (Morales & Trotman, 2004).

Given that an academically resilient student is likely to have several of these attributes, these protective factors tend to influence each other. For instance, a student who is academically motivated will likely have a stronger internal locus of control and a better self-concept, and thus engage in academic planning behaviours, such as goal-setting. Therefore, these definitions are meant to serve as conceptualisations of these constructs rather than strict boundaries separating one protective factor from another.

Related attributes were grouped under these themes. For example, mentorship, role models, or peer social cohesion would fit under 'social support' protective factors or self-concept and internal locus of control would be encompassed by 'cognitive style.' All of the aforementioned psychosocial protective factors were theoretically sound and supported by empirical evidence across studies, and such as deemed acceptable to inform the evidence base of included studies for both reviews (see table 1 for data sources). It is also important

to note that this is not meant to be an exhaustive list of psychosocial protective factors that may improve academic achievement. As mentioned in section 1.3.2, academic resilience is context-specific. As such, the numerous traits and behaviours that a student exhibits may be the direct result of his or her unique social ecology, resulting in assets that may benefit one student but perhaps not another. Therefore, the review author attempted to identify the most prominent, and rigorously-evaluated, academic-related protective factors.

Psychosocial protective Factor	***Subset***	**Citations**
Academic motivation		(Pajares, 2001); (Morales & Trotman, 2004)
Academic self-efficacy/self-regulation	Self-regulated learning	(Bandura, 1997); (Martin & Marsh, 2006); (Zimmerman & Schunk, 2001); (Nota et al., 2004); (Schunk & Zimmerman, 1998)
School-based social support	Peer networks; mentoring; school connectedness	(Johnson et al., 2015); (Morales & Trotman, 2004)
Academic Perseverance		(Martin & Marsh, 2006)
Academic Planning	goal orientation; goal-setting	(Martin & Marsh, 2006); (Jowkar et al., 2014)
Adaptive help-seeking		(Cassidy, 2016)
Coping		(Leary & DeRosier, 2012)
Self-care Behaviours		(Leary & DeRosier, 2012)
Cognitive style (optimism, confidence in one's abilities, etc.)	self-concept; self-esteem; internal locus of control	(Leary & DeRosier, 2012), (Gordon, 1995, 1996); (Strand & Peacock, 2002); (Wang, Haertel, & Walberg, 1994); (Padrón, Waxman, Brown, & Powers, 2000)

Table 1. Academic-related psychosocial protective factors and their relevant evidence base.

1.3.4 Contextualising the use of protective factors in school-based interventions

Garmezy's (1991) 'characteristic triad' model provides a conceptual framework in which protective factors can be contextualised. Grouping factors by 'family,' 'disposition,' and 'environment' allows for a better understanding of the areas where these assets are developed, and thus how to create school-based programmes (see figure 3). Though the triad model is useful in structuring interventions for academic resilience, it is important to

note that certain protective factors may intersect with other domains, and as such cannot be put in discrete categories (Morales & Trotman, 2004). Given the focus on school-based interventions and the particular importance of the school climate in resilience research, this review will focus, where possible, on the domains of environmental and dispositional protective factors which can be fostered in school and exclude interventions in family settings (Garmezy, 1991). Garmezy's model also illustrates the significance of equifinality, which states that the same outcome can arise from numerous processes and factors (Liddle, 1994). Within the context of academic resilience programmes, this means that various types of protective factors can all lead to the same result of instilling a student with positive adaptive behaviours. Therefore, equifinality additionally highlights the multiple, context-specific pathways by which risk factors may surface thus compromising academic performance, and the comparable ways in which protective factors for academic resilience can be implemented to improve educational outcomes.

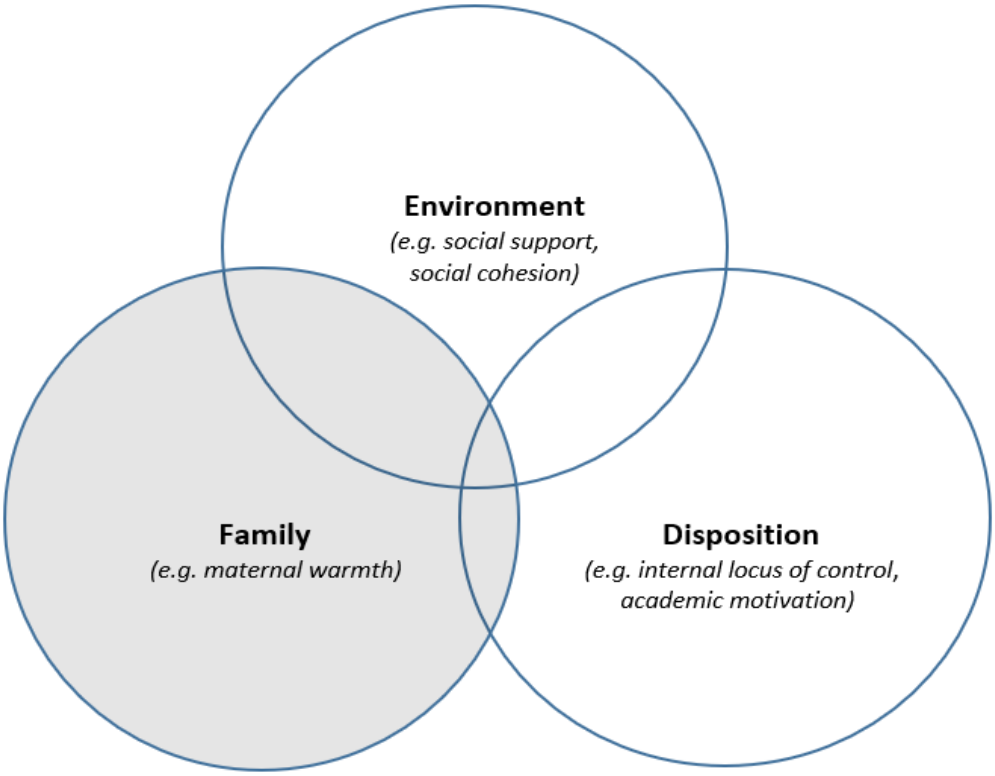


Figure 3. Garmezy's (1991) Characteristic Triad Model with contextual examples of protective factors related to academic resilience. NB: The original model has been modified with the addition to overlapping circles to highlight the interplay between the three domains, and shading in the family domain to signify its exclusion from this review.

Environmental psychosocial protective factors typically use an outside force, such as a group of friends or a positive school authority figure, as a catalyst to foster academic resilience. Pinpointing the environmental factors within the collection of academic protective factors mentioned above, that would include academic motivation, adaptive help-seeking, and especially school-based social support. In terms of environmental (i.e. 'external') psychosocial protective factors, the school community, as previously mentioned, plays a major role in fostering academic resilience as it creates social cohesion and support within a student's learning environment.

Connecting this point to specific interventions, school-based programmes can utilise highly structured classrooms through which students can foster better relationships with their teachers, which have been found to be effective (Morales & Trotman, 2004).

Other effective interventions for promoting educational resilience include the use of peer social support such as established peer social networks (Hartley, 2011; Swenson, Nordstrom, & Hiester, 2008). These interventions utilise either reciprocal student pairs or groups to build social cohesion amongst themselves or assemble small groups of students who create informal, positive relationships with a target student in need, such as the intervention used in Asmus and colleagues (2017) (see table 7 in section 4.1.2 for description of included studies). Additionally, social support in the form of mentoring has been discussed extensively as a salient approach to making at-risk students resilient (Alva, 1991; Gordon & Song, 1994; Villanueva, 1996; Werner & Smith, 1992). These types of interventions allow mentors to fill a vital role in lives of at-risk students by creating intergenerational support due to the students' lack of social capital (e.g. undereducated parents, lack of academic support, etc.) which bolsters their academic endeavours (Villanueva, 1996). Through mentorship, students are given a self-confidence through their mentor's personal investments in them as well as expectations and standards surrounding their academic performance (Levine & Nidiffer, 1996).

Dispositional factors refer to individual behaviours and attitudes that can be augmented to improve school performance instead of individual traits that are unchangeable. Within the scope of this review, relevant dispositional factors are academic self-efficacy and self-regulation, academic perseverance, coping, academic planning, self-care behaviours, and cognitive style. Though dispositional (i.e. 'internal') factors can be fostered within educational surroundings, these types of skills have been differentiated from environmental factors as they similarly increase academic resilience but through different means, by directly targeting certain behavioural and personality characteristics (Morales & Trotman, 2004). In this way, dispositional factors do not require an outside agent in the way that environmental factors do. Referring to cognitive style, having an

internal 'locus of control' is an often-cited psychosocial protective factor (Borman & Overman, 2004; Cappella & Weinstein, 2001; Luthar, 1991). An internal locus of control dictates students' beliefs about their influence over their educational environment and its outcomes. An external locus of control, on the other hand, can lead to learned helplessness, guiding students to believe that they have no agency over their situations (Seligman, 1972; Testa-Ryan, 2016). Perhaps the most pertinent explanation for the utility of an internal locus of control is that for many disadvantaged students who overcome adversity, they are usually the first in their families to do so. Thus, as the research from Morales and Trotman (2004) explains, these students may learn to rely solely on themselves on the path to academic success.

Related to internal loci of control are the concepts of self-efficacy, self-esteem, and self-concept. Tying in specific interventions, Curran (2018) uses mental flexibility training to improve students' self-concepts regarding their mathematical abilities. By simply assigning classrooms by maths competency, the intervention aimed to improve students' internal loci of control by placing in proximity to students who performed similarly, thereby increasing their positive feelings about their own performance. Thus, a student possessing self-efficacy typically feels that his or her hard work will result in academic achievements, which in turn may influence their self-esteem and the belief in their ability to reach certain scholastic goals (i.e. self-concept), all of which have been shown to correlate with heightened academic performance (Cappella & Weinstein, 2001; Gordon, 1995; Luthar, 1991). Given the contextual nature of academic resilience, it is impossible to account for all psychosocial protective factors; however, both the environmental and dispositional assets mentioned above serve as examples of some ways to promote academic resilience.

Targeting interventions at a student's disposition or school environment could disrupt the snowballing effect of compounded risks factors, reduce the severity of one risk that would otherwise cascade into another ecological level, or aim to improve functioning in one domain thereby increasing the likelihood of functioning in other domains (e.g. improving an individual's academic self-regulation could lead to overall improvement in school) (Cicchetti & Curtis, 2007; Masten, Long, Kuo, McCormick, & Desjardins, 2009). In modern-day society where having an education is so important that it is a major predictor of success in later life, it stands to reason that academic risk needs to be minimised as much as possible, and an effective way of tackling adversities is through fostering academic resilience (Austin, 1996; Jackson & Moore, 2006).

1.4 How the intervention might work

School-based interventions utilising psychosocial protective factors as the main components to promote academic resilience have a theorised causal pathway that culminates in improved educational outcomes. Using Johnson and colleagues' (2015) framework for social support via words of encouragement and role models as an example, they assert that these psychosocial protective factors increase the use of regulatory strategy skills (Johnson et al., 2015). Environmental assets such as these are believed to interact with the student's behavioural and cognitive patterns in what is referred to as the 'reciprocal determinism' model (Bandura, Freeman, Mahoney, DeVito, & Martin, 2004). Hence the psychosocial protective factors from the intervention in conjunction with the pupil's individual attributes should precipitate the use of metacognitive regulation. It is hypothesised that students should be more capable of persevering through challenging school experiences (i.e. academically resilient) by using regulatory strategies (time management, self- and effort-regulation), resulting in higher educational performance. The

rationale behind this causal model has been supported by a corpus of literature which argues that resilience and regulatory strategies are related (Dishion & Connell, 2006; Mornane, 2009; Nota et al., 2004), and that these strategies may ameliorate academic difficulties (Duncan & McKeachie, 2005), in addition to the significant findings in the study itself (Johnson et al., 2015). Thus, it is assumed that within these interventions, academic resilience mediates the outcome of scholastic performance through the use of adaptive regulatory skills stemming from protective factors (see figure 5 below).

The impact of the intervention should not end with the programme itself. Rather, it should allow students to continue relying on their academic resilience throughout their educational careers. Modifying Morales' (2000) Resilience Cycle (Morales, 2000) to reflect the above causal pathway, school-based interventions should help at-risk students to:

1. Effectively identify their major psychosocial risk factors.
2. Gain psychosocial protective factors which lead to the development of academic resilience, resulting in necessary adaptive regulatory skills.
3. Use these positive strategies to ameliorate their academic trajectories.
4. Recognise the value of these skills and continue to implement and refine them to address context-specific academic challenges.
5. Sustain their cognitive-behavioural patterns supporting their newfound academic achievement until their desired goal has been reached.

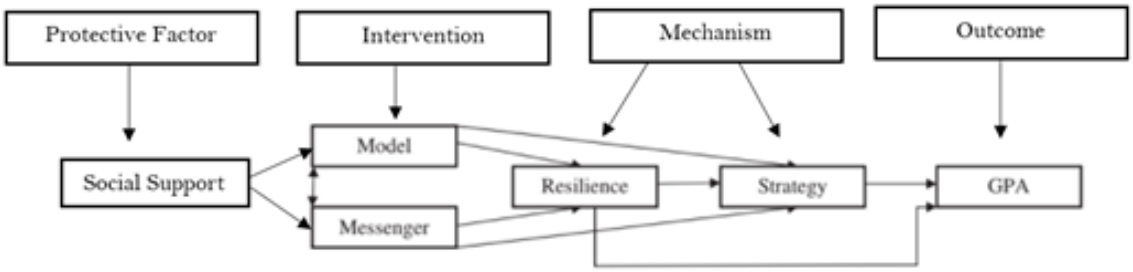


Figure 4. Logic model of intervention outlining the causal impact of protective factors (i.e. models and messengers), the causal mechanisms of academic resilience and regulatory strategies, and the mediated outcome of educational performance (i.e. GPA) (Source: Johnson et al, 2015).

1.5 Differentiating academic resilience and social emotional learning in school-based interventions

To better conceptualise the nature of academic resilience in school-based interventions, it is important to discuss where this concept stands in relation to other types of psychosocial interventions which may affect academic achievement. Arguably social and emotional learning interventions are the most akin to academic resilience interventions, and as such, some similarity exists between the two fields in the research literature. Social and emotional learning has been defined as the process through which children learn and effectively apply the necessary skills to address their emotions, realise positive goals, exhibit compassion and empathy for others, foster positive relationships, and make responsible choices (CASEL, n.d.). Stemming from this definition, the Collaborative for Academic, Social, and Emotional Learning (CASEL) have outlined the following five competencies: self-awareness, self-management, social awareness, relationship skills, and responsible decision-making (CASEL, n.d.).

In terms of the likeness in the research literature, academic resilience and social and emotional learning resemble one another on the SEL competencies of self-awareness, self-management, and relationship skills. Self-awareness encompasses the ability to regulate

one's thoughts and behaviours, and in this way, it is comparable to the protective factor of self-regulation cited in the academic resilience literature (Hu et al., 2011; Nota et al., 2004; Zimmerman & Schunk, 2001). Additionally, self-management, which involves self-discipline, self-motivation, and goal-setting, parallels academic resilience with respect to protective factors like self-efficacy (i.e. not only regulating oneself but putting these behaviours into action), academic motivation, and academic planning, respectively (Martin & Marsh, 2006; Morales & Trotman, 2004; Pajares, 2001). Also similar to the realm of academic resilience is the SEL competency of relationship skills, as social support is an often-cited environmental attribute that improves the likelihood of academic success (Johnson et al., 2015; Morales & Trotman, 2004).

Though the two fields mirror one another to a certain extent, they diverge on the social and emotional learning domains of social awareness and responsible decision-making. Social awareness represents the skills of perspective-taking, empathy, appreciating diversity, and respect for others. Within the scope of the literature on academic resilience as well as the practical use of its interventions, these psychosocial tools are seldom, if ever, used. That is not to say, however, that these attributes could not contribute to a student's *global* resilience, which in turn improves several aspects of life functioning, including academic performance. But within the realm of *academic resilience*, these aforementioned factors are not considered as salient protective factors that would have a direct impact on a student's scholastic abilities. Furthermore, academic resilience differs slightly from social and emotional learning in terms of the fifth SEL competency of responsible decision-making. Specifically, the emphasis on ethical responsibility has not been underscored as an important protective factor as it relates to academic resilience. Since the concepts differ from academic resilience, interventions that would focus on these aforementioned skills would be considered outside of the scope of this review.

A major distinction between the two fields in terms of intervention development and application is the presence of emotional management and conflict resolution programmes. An example of this can be found in the 4Rs programme, a well-known branded SEL intervention that addresses in large part literacy development and conflict resolution. According to the programme's theoretical framework, the intervention provides a 'pedagogical link' between the teachings of conflict resolution and academic content (Brown, Jones, LaRusso, & Aber, 2010, p. 156). This is important to note as it connects to the earlier point about the difference between the two research fields on the domain of social awareness, and specifically respect for others. This programme's attention to fostering respect through conflict resolution would not constitute as something typically found in academic resilience. Given the significance of mutual respect as a facet of SEL programmes, interventions like 4Rs would not apply to academic resilience given the lack of research on conflict resolution as a prominent academic-related protective factor.

A final point of differentiation between the two fields is the outcome objective. The goal of academic resilience is to increase a student's subjective well-being insofar as these psychosocial improvements can directly correlate to academic success. Social and emotional learning, on the other hand, views subjective well-being not as a mechanism to stimulate scholastic performance, but an outcome in and of itself (CASEL, 2012). This then leads to another key difference between the two concepts in terms of the settings in which these outcomes can be used. As noted earlier in this introduction, academic resilience is defined as a context-specific subset of global resilience that functions in the sphere of education. With this thought in mind, academic resilience is nurtured, in large part, in educational environments because the intended outcome is to make students more academically adept. Therefore, it appears that school settings would be the most appropriate place to use academic-related protective factors. Social and emotional learning,

though typically fostered through school-based interventions, provides students with assets that can be applied in a myriad of settings given the added focus on behavioural skills training, which can be used outside of the classroom. Thus, unlike academic resilience, the sole objective of SEL interventions is not academic achievement, though that is one of the several outputs that can stem from these types of programmes.

These distinctions are not to undercut the value of social and emotional learning but rather to simply illustrate that there are multiple ways in which psychosocial programmes can drive educational performance. Academic resilience is only one of several mechanisms that can ameliorate school outcomes. However, within the scope of this review, it is vital to establish clear demarcations between interventions of academic resilience and other similar disciplines to justify the rationale behind the inclusion, and by extension exclusion, of certain studies. Therefore, where feasible, the effort was made to identify and distinguish academic resilience interventions as best as possible. Reports that were considered by the primary study researchers as social and emotional learning were included if they were highly relevant to one of the noted protective factors that promote academic resilience.

1.6 Applying an equity lens

1.6.1 The nature of equity

The terms 'inequality' and 'inequity' have been used in a variety of ways throughout the research literature (Leon, Walt, & Gilson, 2001). Inequality is defined as measurable differences in experience and outcomes across various populations, often dictated by demographics such as location, SES, age, disability, gender and/or ethnic group (Whitehead, 1992). This differs slightly from inequity, which points to differences in

opportunity among groups which can cause unequal life experiences which are deemed unjust (Whitehead, 1992). Modifying the definition from the World Health Organisation's Commission on Social Determinants of Health for this review, inequity can be seen as 'the [presence] of unfair and avoidable or remediable differences in *education* among social groups' (WHO, 2007; p. 9) [emphasis mine]. Though commonly referred to as educational *inequality* throughout the research, the term *inequity* will be prioritised throughout the remainder of this review, as it emphasises the unfair, systemic barriers preventing at-risk students from succeeding academically. When applying an equity lens, it is important to consider if the aim of equity is to reduce educational gradients and/or educational gaps. The educational gap approach focuses on the absolute difference in academic performance between the least and most disadvantaged; whereas educational gradients capture the whole population spectrum—students who are the worst off, those are doing well, and everyone in between (Tugwell et al., 2010). When tackling educational inequity, it is important to address both gaps and gradients where possible to highlight both absolute and relative academic improvement.

1.6.2 PROGRESS-Plus framework

This equity review is informed by the PROGRESS-Plus Framework, a tool developed by the Campbell and Cochrane Collaborations to incorporate equity into review methodology. The acronym **PROGRESS** refers to **P**lace of residence, **R**ace/ethnicity/culture/language, **O**ccupation, **G**ender/sex, **R**eligion, **E**ducation, **S**ocioeconomic status (SES), and **S**ocial Capital (see table 2 below). The framework was expanded to include 'Plus' factors representing additional characteristics such as disability and age. Overall, these factors pinpoint areas in which inequity can occur and as such, highlight the social determinants of education which can be targeted through interventions.

Equity considerations are usually limited to a single category (O'Neill et al., 2014). However, this review will be analysing equity effects across race, gender, and socioeconomic status, the theoretical justifications for which have outlined above (see section 1.2, Description of the Condition).

PROGRESS Plus Factor	Educational Inequity Context
Place of Residence	Place of Residence could negatively impact catchment areas, and thus the quality of local schools.
Race/Ethnicity/Culture/Language	Minority students are more likely to be raised poor, which creates educational barriers.
Occupation	A parent's occupation, or lack thereof, can lead to a low-income household, which research has shown can cause academic underperformance.
Gender/Sex	Stereotype threat has been shown to cause academic difficulties among African-American boys in general and girls specifically in mathematics.
Religion	Strict religious practices and beliefs may preclude access to education in developing contexts.
Education	Low levels of maternal education have been shown to have a negative impact on home literacy rates and academic support, both of which produce academic risk in children.
Socioeconomic Status (SES)	Low SES is one of the major characteristics of educational inequity due to limiting social mobility to access vital academic resources.
Social Capital	Lack of social support is a major risk factor for vulnerable children who lack guidance and understanding of how to navigate challenging academic situations.
Plus – e.g. age, disability, etc.	Physical disabilities and health problems may result in a student being placed in special education, thus making it more difficult to reach educational milestones.

Table 2. Examples of educational inequity using the PROGRESS-Plus framework.

Equity reviews are powerful tools for several reasons. Firstly, they showcase the distribution of effect through the use of harvest plots which summarise intervention effectiveness across social determinants (Humphreys & Ogilvie, 2013). Focusing on the main effects of an intervention can limit its full utility as it does not account for the effects on different types of clients. Related to this point are the concepts of equity harm and intervention-generated inequalities, because not acknowledging the importance of equity in interventions and systematic reviews may in fact widen the gap of inequity, making those most in need worse off (Lorenc & Oliver, 2014; O'Neill et al., 2014). Secondly, they allow for better comprehension of the condition of interest by incorporating additional historical and contemporary information that synthesises the contextual factors which justify the severity of educational inequity, and thus the importance of this review (Tugwell et al., 2010). Thirdly, equity reviews identity gaps in treatment populations, as research shows that programmes typically do not reach the poorest in society (Gwatkin, 2003). Thus, equity reviews provide salient, worthwhile information about the comprehensive effects of interventions, and the use of its methodology has been modified from its epidemiological origins for this purpose of this review.

1.7 Importance of the review

To date, no such review on both the equity and effectiveness of school-based interventions for academic resilience in educational performance exists. Therefore, this review is important in terms of what it can add to the research literature as well as evidence-based policy and practice. In regards to research endeavours, this review is useful in two ways. There is a large area of research dedicated to academic resilience; however, no researchers have synthesised these findings in a systematic way to highlight what might be the most effective (or inversely, futile or potentially harmful) tools in promoting academic

resilience. As such, this review could identify potentially effective intervention types or elucidate gaps in the evidence base where more research is needed. Moreover, given that the equity portion of this review uses the PROGRESS-Plus framework, it introduces a relatively new methodology to both the realm of systematic reviews of interventions and the research field of academic resilience. Currently, most Cochrane equity reviews focus on public health outcomes, so this review aims to extend the use of equity methods to assess educational inequality (Campbell and Cochrane Equity Methods Group, n.d.). Equity reviews should be considered as especially important within academic resilience research, which aims to combat the detrimental effects of educational inequity on vulnerable children. Therefore, this review may be a vital contribution to the literature.

This review could also serve as a guideline for what constitutes as best practice in policy and practice decisions. For practitioners, this review could influence the development and implementation of specific school-based programmes targeted at specific risk factors of vulnerable students. For policymakers, this equity review could inform decisions on the local and/or national level on how to assist schools in getting their at-risk students to perform at grade equivalence or higher. Beyond the utility for researchers, practitioners, and policy makers, this review is important in terms of its ethical considerations. The use of equity in a systematic review underscores the responsibility to ensure the most vulnerable receive proportionate benefit from the intervention by reducing gradients of inequity as much as possible (Tugwell, 2006; Welch, Tugwell, & Morris, 2008). Considering the implications of academic failure on adulthood, including financial insecurity, mental illness, and physical health problems, there is an ethical imperative to ensure that at-risk students perform the best they can through the use of academic resilience (Caspi, Wright, Moffitt, & Silva, 1998; Felliti & Anda, 1997).

2 Objectives and research questions

Given the interest in both effectiveness and equity, this review has two research questions and objectives, guided by the PROGRESS-framework shown above as well as additional guidance from the PICOS framework (see table 3 below).

2.1 Effectiveness

With respect to the effectiveness section of this review, the research question is the following: **Are school-based psychosocial interventions which target academic-related protective factors compared to treatment as usual effective in improving educational performance among primary and secondary students?**

The research objective here to assess the effectiveness of school-based interventions fostering academic resilience on academic gains among primary and secondary students.

2.2 Equity

The objective of the equity section of this review is to determine if school-based interventions produce equitable outcomes in academic performance among students who are at higher risk due to certain characteristics (i.e. race, gender, and SES) relative to their lower-risk peers. The research question informing this portion of the review is the following: **Do school-based psychosocial interventions fostering academic resilience have equitable effects on academic gains across factors of race, gender, and socioeconomic status among primary and secondary students?**

Concept	Definition
P - Population	Primary and secondary students
I - Intervention	Psychosocial interventions that promote protective factors related to academic resilience, e.g. social support, academic motivation, self-regulation, etc.
C - Comparison	Compared to treatment as usual (e.g. normal school curriculum) or other intervention types (e.g. academic tutoring)
O - Outcomes	Improved academic performance, measured through GPA, attendance, school conduct, school retention, and educational milestones (e.g. graduating, moving on to the next grade, etc.)
S - Setting	Within the school environment during normal school hours

Table 3. PICOS Framework.

3 Methods

3.1 Methodology

3.1.1 Rationale for selection criteria

3.1.1.1 Study design

Both experimental and quasi-experimental study designs were considered for inclusion. Though randomised controlled trials are more able to determine causality and produce unbiased evidence over other design types, potential ethical concerns and additional restrictions, such as school policies, can prohibit randomisation of students in educational research (Baron, Evangelou, Malmberg, & Melendez-Torres, 2017). Thus, this review included quasi-experimental designs as long as the studies assessed intervention effect and attempted to control for potential confounders. The following quasi-

experimental designs were eligible for inclusion: natural experiments, pretest-posttest designs with control, regression discontinuity design, matched control group designs, and interrupted time series designs. It is worth noting that certain non-randomised designs are weaker than others, such as pretest posttest with control; however, the literature on equity reviews states that a broader range of study types should be considered given their fitness for purpose at addressing research questions of interest (O'Neill et al., 2014; Tugwell et al., 2010; Vivian Welch et al., 2012, 2016). Thus, methodological appropriateness has been prioritised over the typical evidence hierarchy (Petticrew & Roberts, 2003), which may have otherwise excluded quasi-experimental designs, which would have limited the evidence base of this review. Qualitative designs were excluded from this review because qualitative studies do not assess causal intervention effects.

3.1.1.2 Study participants

The study population was limited to primary and secondary school students as the most deleterious effects of academic underperformance and educational inequality occur at this point in schooling (Banerjee, 2017; Eckenrode, Laird, & Doris, 1993; Leiter & Johnsen, 1994). This timeframe within a student's education may be a very critical point for intervention, preventing these academic shortcomings from either dampening a student's performance in university or hindering students from pursuing a post-secondary education entirely. Particularly, the transition from elementary school to high school has been known to cause unwanted changes for students in term of school resources and social support networks, in addition to increasing academic difficulty, all of which may cause academic failure (Hanewald, 2013; Martínez, Aricak, Graves, Peters-Myszak, & Nellis, 2011). Primary and secondary students of varying racial, gender, and socioeconomic backgrounds were

included in this review in order to assess the distribution of equity effects across the social determinants of education.

It was surmised that kindergarten and pre-school students would not be appropriate to include the review study population as early childhood may be too young of an age to clearly see the effects of academic underperformance since they are rarely formally assessed at this age. Instead, there is more of an emphasis on school conduct and behaviour, which are important prerequisites of school functioning, but do not necessarily equate to academic success. Post-secondary students were also excluded from this review because, as alluded to above, it may be too late to affect their academic trajectory in such a way that can have a significant change in their educational performance that would lead to better life outcomes, like improved employability based on their educational status.

3.1.1.3 Intervention, settings, and outcomes

Intervention components eligible for inclusion target at least one of the listed psychosocial protective factors. Some school-based interventions may contain non-psychosocial aspects as well (i.e. academic, such as reading and math), so studies combined with another programme or intervention type were included if 50% or more of the intervention components related to the aforementioned protective factor(s). In terms of settings, out-of-school interventions, such as summer school or afterschool programmes, were considered beyond the scope of this review, and for this reason, they were not eligible for inclusion. Additionally, there was no geographical restriction on eligibility. Given the focus on improving scholastic abilities, the only outcome of interest is academic achievement. The academic outcomes were ranked in order of importance as certain outcomes are more likely to be reported and are better measures of school performance, such as school grades. As per the Cochrane standards for systematic reviews of

interventions, studies that did not readily report outcome data on academic achievement were not excluded. All authors of eligible studies without academic outcome data were contacted, if no other reason for exclusion was found. All of the academic achievement outcome measures listed below were eligible for inclusion since there are many ways to assess academic performance.

3.2 Criteria for considering studies for this review

3.2.1 Types of studies

Studies with experimental or quasi-experimental designs that assess treatment effect were considered for inclusion in this review. Though some designs are stronger than others, they were ranked in no particular order. The following eligible studies included:

a. Randomised controlled trials: Random assignment of participants to either treatment or control groups by the researchers, using an unbiased method of random allocation (e.g. random number generator)
b. Natural experiments: Assignment to control versus treatment group in which allocation is determined by nature rather than the researcher, though it may invariably mimic random assignment which allows for causal inference.
c. Pretest-posttest designs with control: Non-randomly assigned treatment and control groups, that are assessed pre- and post-intervention. The pretest accounts for potential baseline differences between the groups, and posttest pinpoints any measurable differences which may be attributable to the intervention.

d. Regression discontinuity design: A threshold is assigned (e.g. a certain percentile on an aptitude test), above or below which the intervention is delivered. Though randomisation does not occur, thus increasing the likelihood of bias, the participants' results that are near either side of the cut-off allow for an estimation of a treatment effect.
e. Matched control group designs (incl. propensity score matching): Treatment group is compared to a matched group of control participants, who are similar on one or more a priori characteristics but do not receive the intervention.
f. Interrupted time series designs: Participants are assessed at multiple time points (i.e. before, during, and after) to determine whether changes are attributable to trends over time or intervention effectiveness.

3.2.2 Types of participants

Studies were included if the participants were students attending primary or secondary school.

3.2.3 Types of intervention

Eligible studies analysed the effect of psychosocial interventions that target specific psychosocial protective factors known to improve academic performance (see below) in comparison to treatment-as-usual curricula or other intervention types. Treatment-as-usual would be what the school used pre-intervention, such as the school's normal curriculum. Other intervention types were those introduced concurrently with the psychoeducational programmes as a means of comparison,

such as academic tutoring. Studies were included if at least 50% of the intervention directly related to one or more of the academic protective factors.

Evidence-based academic protective factors:

a. Academic motivation: A student's interest in being academically inclined given the perceived benefits of educational achievement.
b. Academic self-efficacy and self-regulation (incl. self-regulated learning): The combined beliefs and behaviours that allow a student to utilise and manage their own scholastic abilities.
c. Coping: Various personal, social, and behavioural abilities that affect how students comfort themselves in stressful academic situations.
d. School-based social support: The level of meaningful involvement and interaction that a student has with his or her school environment.
e. Academic perseverance: The sheer determination to persist through academically challenging experiences, which may include prolonged exposure to periods of chronic stress.
f. Academic Planning: Goal-orientation and goal-setting strategies used to effectively accomplish a specific educational endeavour.
g. Adaptive help-seeking: The process by which students to identify their own learning obstacles and then actively seek out assistance from others who can help them.
h. Self-care behaviours: The use of health-promoting behaviours during times of stress which in turn improve a student's functioning.
i. Cognitive style (self-concept, internal locus of control, etc.): The thoughts, beliefs, and attitudes through which students attribute value to their own abilities and their outlook on their academic future.

3.2.4 Types of outcome measures

3.2.4.1 Primary outcomes

Various measures of academic performance were included in this review. In line with Cochrane standards, studies were included whether or not they reported academic outcomes. The authors of studies who did not report academic outcome data were contacted. The following outcomes were ranked in order of importance:

a. School grades, i.e. grade point average (GPA) or specific course marks
b. Academic test scores (excl. IQ tests, vocal tests, etc.)
c. School attendance
d. Grade level progression
e. Educational milestones (graduation, etc.)
f. School conduct and behaviour

3.2.4.2 Secondary outcomes

There were no secondary outcomes reported in this review.

3.2.5 Duration of follow-up

Data from any duration of follow-up in the original studies were included in the review. Follow-up data were classified by short term (0-5 months post-intervention), medium term (6-11 months post-intervention), and long term (12 months or longer post-intervention).

3.2.6 Types of settings

Given the review's focus on school-based interventions, only studies delivered within school during normal school hours were included.

3.2.7 Types of language

Only studies reported in English were eligible for inclusion.

3.2.8 Reasons for exclusion

As studies may have multiple reasons for exclusion, an exclusion hierarchy was applied. Firstly, studies were considered ineligible if they did not contain an intervention component of interest or did not meet the 50% threshold if the study contained additional components not listed above. Secondly, excluded studies used at least 50% or more of a study population different from primary and secondary students, such as a study based on solely undergraduates. Thirdly, studies outside of the scope of the setting, including non-school locations and before-, after-, and summer-school programmes, were excluded. Any additional reasons for exclusion were weighted equally.

3.3 Search methods for identification of studies

3.3.1 Electronic searches

3.3.1.1 Database selection

Sixty-one data sources were piloted to identify any mention of relevant information in the title, abstract, or text body. A combination of key words, such as 'intervention' and 'protective factor' and 'academic achievement,' were searched or were replaced with similar terms depending on the subject indexing of each data source. From this process, only six electronic databases produced potentially relevant hits. The full list of databases searched is as followed:

Database	Search Date (from database inception unless otherwise stated)	Number of Hits
1. Applied Social Sciences Index & Abstract (ASSIA)	30 December 2017	140
2. Embase (Ovid)	1974 - 19 November 2017	1885
3. Education Resources Information Center (ERIC)	30 December 2017	43
4. MEDLINE (Ovid)	1946 - 18 November 2017	945
5. ProQuest Dissertations & Theses Database (full Global Text)	27 December 2017	135
6. PsycINFO	1967 - 30 December 2017	1465

Table 4. List of electronic databases

3.3.2 Search string

A comprehensive, highly sensitive search string was developed iteratively to encompass the full research scope of the review given the conceptual complexity of academic resilience. Search terms were initially informed by the research literature on academic resilience (see section 1.3.3), and later augmented after piloting based on database subject mapping and reading the abstracts of potentially relevant texts. To further increase sensitivity, key search terms were expanded upon using truncations and wildcards, as well as synonyms from www.thesaurus.com. Table 5 below provides an example of a search string used in MEDLINE.

Research Component	#	Search History	Results
P O P U	1	Child/	1660206
	2	Adolescent/	1982161
	3	Students/	50385
	4	Child*.mp.	2357192
	5	Adolescen*.mp.	2044967
	6	Student*.mp.	278330
	7	Schoolchild*.mp.	13182
	8	Elementary school*.mp.	9102
	9	Primary school*.mp.	10947

LATION	10	Middle school*.mp.	5032
	11	Junior high*.mp.	2393
	12	High school*.mp.	28654
	13	Secondary school*.mp.	9318
	14	(1st grade* or 2nd grade* or 3rd grade* or 4th grade* or 5th grade* or 6th grade* or 7th grade* or 8th grade* or 9th grade* or 10th grade* or 11th grade* or 12th grade*).mp.	6552
	15	(Grade 1 or Grade 2 or Grade 3 or Grade 4 or Grade 5 or Grade 6 or Grade 7 or Grade 8 or Grade 9 or Grade 10 or Grade 11 or Grade 12).mp.	71582
	16	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15	3661436
	17	((undergrad* or graduat* or dental* or nurs* or medic* or law or business) adj3 student*).mp.	99080
	18	16 not 17	3562356
INTERVENTION	19	Resilience, Psychological/	3752
	20	Adaptation, Psychological/	91849
	21	Grit*.mp.	2754
	22	(academic resilience or academic hardiness or educational hardiness or academic adapt* or educational adapt* or academic persistence or educational resilience or educational persistence or academic buoyanc*).mp.	66
	23	(based or focus?ed ! or (informed adj3 resilience)).mp.	3018397
	24	(academic motivat* or educational motivat*).mp.	220
	25	self efficacy/	17796
	26	Self-Control/ or Internal-External Control/	19635
	27	Coping.mp.	48573
	28	self-regulated learning.mp.	323
	29	Social Support/	66850
	30	persever*.mp.	5006

	31	help?seeking.mp.	46
	32	(cognitive style* or optimis* or pessimis*).mp.	
	33	(factor* adj3 (protect* or promoti* or external or internal or environment*)).mp.	119486
	34	(competenc* adj3 (emotional or behavio?ral or social or cognitive)).mp.	4678
	35	(positive psychology or (psycho adj1 social) or positive education).mp.	4460
	36	(Penn Resilien* Programme* or PERMA or FRIENDS for life or Resilient Families Programme* or Adult Resilien* Programme* or Friends Resilience).mp.	128
	37	19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36	3317345
	38	(school* adj3 (intervention* or programme* or course* or practice* or curricul* or environment*)).ti,ab.	23010
	39	37 and 38	10563
OUTCOME	40	Educational Status/	50082
	41	Educational Measurement/	35170
	42	student dropouts/	1741
	43	(Academic performance or Academic achievement or Educational performance or Educational achievement or Scholastic achievement or Scholastic performance).mp.	10045
	44	(GPA or Grade point average or Cumulative grade point average or Mark or Academic grade* or Standardi*ed test score* or Student assessment* or Academic assessment* or Percentile*).mp.	70366
	45	(Student conduct or Academic readiness or! Attitude* towards school).mp.	66
	46	(Grade retention or School retention or Grade Repetition or Attendance or	38754

		Tard* or Truan* or School refusal).mp.	
	47	(Academic progress* or graduation rate* or School comple*).mp.	884
	48	40 or 41 or 42 or 43 or 44 or 45 or 46 or 47	197242
	49	18 and 39 and 48	945

Table 5. Search String – MEDLINE (Ovid)

3.3.3 Searching other resources

In addition to searching electronic databases, the references of included studies, Google Scholar, and grey literature sites were checked for additional relevant studies. Furthermore, research experts were contacted to locate information about ongoing and unpublished studies (see appendix 7.7).

3.4 Data collection and analysis

3.4.1 Selection of studies

The review author conducted eligibility screening on all studies. Of these texts, a random sample of approximately 20% of eligible studies (n=26) were screened independently and concurrently by a second review (P.D.). Both researchers reviewed the title, abstract, and full text to determine whether the study was suitable for inclusion. This was done to check the comprehensiveness and accuracy of the eligibility criteria given the conceptual complexity of the subject matter. To quantify the level of agreement between the two reviewers, Cohen's kappa was calculated, which was 0.71. Kappa statistics between 0.61 and 0.80 are considered as substantial, so the selection criteria were deemed as an accurate means of identifying the correct studies for this review (Landis & Koch, 1977). Any disagreements about study inclusion and exclusion were resolved through discussion.

3.4.2 Data extraction and management

The included studies were coded using a data extraction form developed *a priori* for this review. The extraction form was informed largely by the Cochrane Collaboration's data collection form for interventions. It was modified to better include quasi-experimental designs as well as the PROGRESS-Plus data needed for the equity portion of this review (see appendix 7.2). The following section details the data extraction process for both effectiveness data and equity data. All data were managed solely by the review author.

3.4.2.1 Effectiveness data extraction

For the purposes of data extraction and analysis, outcome measures were arranged in a hierarchy, and based on what was reported in the study, data from the highest ranked measure was included in the data synthesis process. The hierarchy below ranks the measures from the most objective onwards:

a. School grades, i.e. grade point average (GPA) or specific course marks
b. Test scores (standardised, individualised, etc.)
c. School attendance
d. Grade level progression
e. Educational milestones (graduation, etc.)
f. School conduct and behaviour
g. Other outcome measures not list above

This approach was undertaken for several reasons. First, certain measures of academic achievement determine school success better than others. For instance, grade point average may be a stronger indicator of academic performance than school attendance, because attending school does not necessarily mean that a student is performing well. Second, this approach is a better option than pooling data across

measures into one effect estimate, which may introduce unintended statistical error. Averaging effect sizes does not easily account for data dependency because these effect sizes are from the sample population, and the resulting variance for each effect size is rarely summarised correctly to reflect a smaller, more accurate composite sampling variance. Therefore, ignoring the dependence will increase the probability of a type I error due to the overestimation of effect sizes and the underestimation of standard error (Moeyaert et al., 2017). Third, conducting a multi-level meta-analysis introduces complications as a small number of outcomes can result in unstable and perhaps negative variance at higher levels. Though multi-level meta-analyses do not require multiple outcome measures, additional research is needed on the accuracy of this analysis when using only one effect size (Van den Noortgate, López-López, Marín-Martínez, & Sánchez-Meca, 2015). Though multilevel meta-analyses are preferable as they typically provide more robust findings, many studies in this review report only one academic outcome measure. As such, pursuing these different statistical approaches for this review may not have been prudent.

3.4.2.2 Equity data extraction

This data process was informed in large part by Humphreys' and Ogilvie's (2013) equity review on population-based physical activity interventions, as well as Welch's (2010) Cochrane methodology review on health equity assessments (Humphreys & Ogilvie, 2013; Welch et al., 2010). Additional information was also obtained from equity methodology papers (Tugwell et al., 2010; Welch, Petticrew, Ueffing, et al., 2012). As mentioned above, the Cochrane data extraction form was modified to capture equity information across the PROGRESS-Plus factors of race, gender, and socioeconomic status. A differentiation was made between equity evidence and equity-relevant data as extracting these types of information influence the review in different ways.

Equity evidence. Equity evidence refers to the statistical analysis of data based on at least one PROGRESS-Plus factor. This includes interaction effects, such as the results from analyses of variance, and subgroup analyses for moderator effects. As moderator analyses may produce spurious results, they are often interpreted cautiously given their methodological limitations, such as small sample sizes and inadequate statistical power (Petticrew et al., 2011). However, if conducted judiciously, they can provide the vital data needed to assess whether or not interventions have equitable impact across the social determinants of education. Additionally, interaction effects suggest if intervention effectiveness is contingent on a PROGRESS plus factor of interest, such as female students in a treatment condition. Unlike equity-relevant data, only equity evidence is included in the harvest plots as they directly assess equity effects. However, given the potential issues due to low statistical power or small sample sizes, one should be cautious about interpreting these findings.

Equity-relevant data. Equity-relevant data underscores the presence of PROGRESS-Plus factors, such as race, gender, and socioeconomic status, but are not used in moderator analyses or interaction analyses. This type of data refers specifically to sociodemographic information and adjusted associations. Adjusted associations are outcomes which have been changed to account for covariates that are PROGRESS-Plus factors (Welch et al., 2010). For example, this would include adjusting the results for lower SES status students to account for the effects of low income on educational performance. Though these types of information do not directly influence the analyses that form equity evidence, demographic information is beneficial as it could guide researchers towards said analyses, which may inform the equity considerations of school-based interventions (Welch et al., 2010). Moreover, adjusted associations are usually performed a pre-requisite step before further analyses to assess the impact of interventions, as a way of accounting for

PROGRESS-Plus factors. Therefore, it is worthwhile information in determining equity effects. It is important to note that data from study populations that were homogenous across the social determinant(s) were not extracted for the purposes of this review (i.e. 100% female population). This is because equity impacts cannot be inferred if students come from similar racial, gender, and/or socioeconomic backgrounds.

3.4.3 Data synthesis

3.4.3.1 Effectiveness synthesis

For the effectiveness portion of this review, study data were meta-analysed using a random effects model as this model better accounts for between study variability in effect sizes.

3.4.3.2 Equity synthesis

For the equity portion of this review, a harvest plot was used to graphically summarise the aforementioned equity evidence extracted from included studies, which was designed using an online diagram programme. The harvest plot method was developed to synthesise evidence of intervention effectiveness across the social determinants of education, i.e. race, gender, and socioeconomic status. (Ogilvie et al., 2008). It is based on a hypothesis-testing approach in which a null hypothesis as well as two alternative hypotheses are postulated. A null hypothesis means that there was no intervention effect on any groups for any social determinant. The positive hypothesis suggests that the intervention effect favours advantaged groups, whereas the negative hypothesis suggests greater intervention effectiveness among disadvantaged groups. Studies are then assigned to the hypothesis for which it has the most supporting data using a matrix of equity evidence across PROGRESS-Plus factors (see appendix 7.3). The negative hypothesis is

desirable because it suggests that an intervention may be effective in improving educational performance among those students who are the most risk (Humphreys & Ogilvie, 2013). Therefore, it is useful in reducing educational inequity. Each bar on a harvest plot represents a single study with its citation number inscribed, which is also weighted by the assessment of study quality. Studies were also evaluated by outcome data metrics, be it direct observation, self-report, or objective measures, and design types (i.e. randomised or non-randomised).

In addition to risk of bias ratings detailed in the next section, studies with equity data were also assessed on two methodological quality criteria. The first criterion, suitability of study design, scored included studies by the rigour of their study design, with one being the lowest score, such as a cross sectional study, and four being the highest score, such as an RCT. The second criterion, methodological quality, was used to evaluate the study's methodological features by six domains, each of which equalled one point. Studies were judged based on the following: randomisation, representativeness, comparability of control group (to treatment group), credibility of data collection tools, attrition, and the causal inferences of the findings (attributability). Table 6 below outlines the coding scheme for equity data synthesis, which was adapted from Humphreys' and Ogilvie's (2013), informed by Kahn and colleagues (2002) and Hillsdon and colleauges (2004). Finally, in addition to the overall quality of the study, the review author considered the specific quality and rigour of statistical analyses used to create equity evidence included in the harvest plot.

Feature	Harvest Plot Visualisation	Coding Rubric
Suitability of study design *(Kahn et al, 2002)*	Bar Height (1-4) represented the level of study design suitability to evaluate equity impacts NB: all studies included in the harvest plot were considered as level 4	**Level 1:** Exposure and outcome measured at a single point in time (i.e. cross-sectional) **Level 2:** Single 'before' and 'after' measurements without a concurrent control group (i.e. single-subject) **Level 3:** At least two 'before' and 'after' measurements without a concurrent control group **Level 4:** Prospective measurement with concurrent control group
Methodological quality *(Hillsdon et al, 2004)*	The number above each bar denoting its overall score for methodological execution of the study. Studies are scored in integers from 0 to 6, depending on how many methodological features are present in each study	**Representativeness**: Random sample recruitment (response rate at least 60%) OR otherwise representative of the study population **Randomisation**: Random allocation of participants, groups, or areas to intervention or control conditions **Comparability**: Comparable baseline characteristics of the comparison group(s) OR important potential confounders were appropriately adjusted for in the analysis **Credibility of data collection instruments**: Valid/reliable data collection tools (e.g. published in other research, widely recognized, etc.) **Attrition rate**: 30% or less missing outcome data **Attributability to intervention**: The likelihood that the outcome can be attributed to the effect of the intervention (NB: this criterion cannot be met if there was evidence of contamination of a control)
Educational performance-related outcome measures	Bar Colour (white, grey, black)	**White**: Directly observed (e.g. researchers collecting achievement data under classroom conditions) **Grey**: Self-reported measures (e.g. a student reporting their test scores on a questionnaire) **Black**: Objective measures (e.g. official transcripts from the school district administrative staff)
Supported social gradient hypothesis	Bar position (left, centre, right)	**Left:** Negative social gradient (e.g. students more likely to be affected by educational inequity) **Centre:** No social gradient **Right:** Positive social gradient (e.g. students less likely to be affected by educational inequity)

Table 6. Harvest plot coding scheme, adapted from Humphreys' and Ogilvie's (2013) equity review.

3.4.4 Assessment of risk of bias in included studies

Risk of bias was conducted separately from the data extraction and management process as the risk of bias section on the Cochrane data collection form was quite limited. Instead, the Cochrane Risk of Bias tool 2.0 and the Cochrane Risk of Bias In Non-randomised Studies – of Interventions (ROBINS-I) assessment tool were used for experimental studies and quasi-experimental studies, respectively (Higgins et al., 2016; Sterne et al., 2016). The decision was made to use two separate risk of bias tools because using the RoB 2.0 parameters to assess the study quality of quasi-experiment designs would lead to an increased likelihood of bias due to the lack of random allocation, which is inherent only to randomised controlled trials. It is important to use the right measurements of critical appraisal when judging study quality as this is a key determinant in the validity of a study's findings. Therefore, to appropriately evaluate the level of bias of non-randomised trials, the ROBINS-I tool was used for quasi-experimental studies.

Experimental studies were rated by the following the risk of bias criteria: random sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, and selective reporting. For each criterion, experimental studies were given a risk of bias rating of either low, high, or some concern. Quasi-experimental studies were rated using the following criteria: confounding, participation selection, classification of interventions, deviations from intended interventions, missing data, measurement of outcomes, and selective reporting. Similarly, quasi-experimental studies were rated by criterion as low-, moderate-, serious-, or critical-risk, and if there were insufficient data, it was classified under 'no information.'

The aforementioned domains were used to inform an overall risk of bias rating for each study. If an experimental study received a risk of bias rating other than 'low' for at

least one domain, such as 'some concerns' or 'high', then the study was given that rating overall. For quasi-experimental studies, both low risk and moderate risk of bias were determined if this rating was applied across all domains. The decision of serious or critical risk of bias was chosen overall if a study received either rating on at least one domain. The rating of no information was applied based on the review author's judgement.

In both the RoB 2.0 and the ROBINS-I tools, each criterion was labelled with a corresponding type of bias, except for the domain of deviations from intended interventions. This criterion refers to multiple biases that were already addressed by other criteria and for this reason, it was left unreferenced. Across all study quality assessment domains, equity data considerations were the most salient in weighing the risk of bias for selective reporting. Given that so few studies presented findings for the PROGRESS-Plus factors of interest, the review author was left to determine if authors deliberately decided against reporting these findings or if the equity data analyses were not conducted in the first place.

Visual representations of risk of bias for both randomised and non-randomised studies were generated in the Cochrane review manager programme, RevMan 5.3 (RevMan, 2014). Given the limitations of the software, changes to the risk of bias graphs and tables could not be made to accurately reflect the RoB 2.0 and the ROBINS-I. As such, unclear risk, designated as a yellow question, refers to 'some concerns' or 'moderate risk' and 'no information' for randomised and non-randomised trials. 'Serious risk' and 'critical risk' for quasi-experimental designs were combined and labelled under the heading of 'high risk', which is signified by a red minus symbol.

3.4.5 Measures of treatment effect

For the effectiveness portion of this review, the review author used the standardised mean difference, Hedges' *g*, because it is a more conservative estimate, and thus more accurate for studies with smaller sample sizes ($n < 30$). Any study that reported another type of effect size was then recalculated using Hedges' *g*. In the few cases of multi-arm trials in where there were more than one treatment group, effect sizes were calculated using the control group and the treatment group that best fit the inclusion criteria. For example, if a study included an intervention group whose programme content was aligned with academic protective factors and another intervention arm which combined this with tutoring, the effect size was calculated using the first intervention group compared to the control group as it better represents the type of interventions of interest for this review. Similarly, if a study used two comparison groups, the effect size was calculated using the control arm was the most similar to treatment-as-usual, as this better reflects the impact of the intervention by comparing it to pre-intervention conditions. All meta-analyses and robustness checks were conducted in RStudio (RStudio Team, 2015).

For the equity portion of this review, treatment effect was assessed using the aforementioned measures: moderator effects, interaction effects, adjusted associations, or baseline socio-demographic characteristics on the PROGRESS-Plus factors of race, gender, and socioeconomic status.

3.4.6 Unit of analysis issues

The sample size was the based on the number of studies in this review. In some cases, multiple reports of one study some were needed for data extraction purposes to provide enough evidence for the review's effectiveness and equity analyses. For clustered trials, i.e. those using a unit of allocation at the group level, experimental or quasi-

experimental, effect sizes were calculated using an intraclass correlation coefficient (ICC) as needed. This was done for studies in which students were nested in the same classroom or school, to account for data dependency as students within the same proximity were likely to affect each other's academic performance. Thus, students' outcome data may not be statistically independent from one another, which if not corrected for, would increase the likelihood of a type I error occurring by overinflating the impact of the intervention (Baldwin, Murray, & Shadish, 2005; Baron et al., 2017). ICCs were not reported in most studies, so commonly-cited values were substituted in from literature on ICCs and grouped trials in education research (Hedges & Hedberg, 2007).

These values were chosen from the literature based on the participants' grade level. If there were students from various grade levels, but one grade level appeared to have a greater number of students than the others, then the ICC was chosen based on the grade level with the most students. Otherwise, if balanced across different levels of schooling, then the ICCs for each grade represented were averaged to create a new value that better reflected the study's sample population. Unit of analysis issues had no bearing on distribution of treatment by the PROGRESS-Plus factors of interest. This is because race, gender, and socioeconomic status are considered as characteristics of each student, so they cannot be significantly impacted by the level at which treatment groups are allocated. Therefore, the equity analysis would not be as affected by unit of analysis considerations as analyses for intervention effectiveness.

3.4.7 Dealing with missing data

The authors of included studies were contacted multiple times to locate missing data. Additionally, the review author searched for online files, such as supplementary materials and protocols, and extracted relevant data that were missing from the full text if

additional information was found. In the cases where if missing data vital to forest plots and/or harvest plots were not retrieved, then the study was not used in the data synthesis process but summarised in the narrative synthesis portion of this review.

3.4.8 Assessment of heterogeneity

For the effectiveness portion of this review, it was initially assumed that heterogeneity would not pose a significant problem for meta-analysis in terms of outcome measures and study population. This is because most studies reported a grade-specific outcome measure of academic achievement (i.e. GPA, math or literacy grade, test scores, etc.) which tends to be measured in similar ways. Additionally, given that the inclusion of student populations was limited to those in primary and/or secondary school, the range of student ages and school levels across studies appeared similar enough for a meta-analysis. To control for heterogeneity in term of interventions, the review author initially intended to compare studies according to their theories of change. Though the intervention components are important, what is of greater significance is the programme theory, all of which relate to academic resilience, which is a causal mechanism in and of itself. Therefore, comparing the interventions by function rather than their form not only provides an acceptable level of homogeneity across studies, it also better reflects the nature of programme effectiveness (Durlak, 1998; Hawe, Shiell, & Riley, 2004). To account for heterogeneity statistically, heterogeneity checks were performed using a random effects model to generate I^2, Tau, and Cochran's Q metrics. Additionally, influence diagnostics were conducted to isolate which study contributes the most to the between-study variability.

For the equity portion of this review, some heterogeneity in terms of gender, race, and socioeconomic status was necessary to assess the distribution of the intervention

effect. However, it was determined that the level of heterogeneity needed for the equity synthesis would not have precluded the use of meta-analyses in the effectiveness portion of the review. This was because heterogeneity was being assessed across gender, race, and socioeconomic status, which are common socio-demographic factors reported in studies. As such, it would not have impacted the quality of the studies, and thus any meta-analytic tools used.

3.4.9 Assessment of reporting bias

Reporting bias occurs when the decision to report findings is influenced by whether or not the results support the research hypothesis (Higgins, Altman, & Sterne, 2011). Positive findings in this review referred to studies that included statistically significant results supporting intervention effectiveness and/or at least one PROGRESS-Plus factor of interest. To ensure that inclusion was not limited to only studies with positive findings, the search strategy was designed to be as comprehensive and sensitive as possible to locate all eligible texts irrespective of the direction of results. This included contacting relevant researchers for additional evidence, searching across 61 data sources, and reference checking included studies. Moreover, a funnel plot was generated to evaluate publication bias, which stems from reporting bias as studies without positive findings are less likely to be published, and vice versa.

3.4.10 Subgroup analyses

Both within-trials and between-trials subgroup analyses were considered for this review. The 'within-trials' type meta-analyses data at the participant level across studies. For instance, if several included studies reported moderator analyses on the PROGRESS-Plus factor(s) of interest within their own trials, these data could then be used in the review's

subgroup analyses. This differs from between-trials subgroup analyses, which, as the name implies, synthesise data across trials instead of comparing the results of participants within the trial. Using race as an example, a between-trials moderator analysis would compare the findings at the trial level between studies in which most participants of one study are of a similar race (i.e. Black) versus other studies where the students are mostly one other race (i.e. White, Asian, etc).

Both types have their limitations. Namely, most trials do not report sufficient information for a within-trial subgroup analysis as most studies do not perform their own moderator analyses, let alone report data by race, gender, or SES. Between-trial analyses have the hindrance of deciding what constitutes as 'most' of one race, gender, or SES. Consequently, this approach must ignore within-study variability across these characteristics. This is an issue because it introduces the added problem of including people who may not fit under of the chosen categorisation since they cannot be parsed out, the inclusion of their data could bias the overall findings. The review author had originally planned to conduct between-trial analyses for each of the PROGRESS-Plus factors as it was assumed that there would be insufficient data for within-trial analyses.

Ideally, researchers could conduct an individual participant data (IPD) meta-analysis. This type of analysis solves these aforementioned problems because it synthesises data at the participant-level across all trials. However, IPD meta-analyses can be costly and time-consuming. Accumulating data for each participant from an included study in a review could take months, if not years, especially if there are many trials. Given the review author's practical limitations, IPD meta-analyses were not feasible for this review but rather noted as a worthwhile solution to the aforementioned methodological limitations.

3.4.11 Sensitivity analysis

Sensitivity analyses were conducted according to risk of bias ratings to determine if study quality impacted the quality of the meta-analysis for main effects. This included both experimental and quasi-experimental studies with a low risk of bias rating. Unlike moderator analyses, sensitivity analyses were not applicable to the equity portion of this review as they assess the quality of evidence rather than its effect across PROGRESS-Plus factors.

4 Results

This section will present the results from the search, a description of included studies, as well as findings for the research questions on both effectiveness and equity. Additionally, this section will outline the risk of bias for randomised and non-randomised studies and the robustness checks to investigate issues with heterogeneity and study quality.

4.1 Description of studies

4.1.1 Results of the search

The electronic database search, grey literature search, reference-checking and contacting authors for relevant studies yielded a total of 4,642 hits (see figure 6 for PRISMA diagram). A majority of these abstracts, 4,295 hits, were identified through OVID (i.e. EMBASE, PsycINFO, and MEDLINE). 318 hits were retrieved through ProQuest (ASSIA, Eric, and the ProQuest Global Dissertation and Theses Database), and 29 additional records were found through other sources, including smaller data sources and texts from authors in the field of academic resilience. After removing duplicates (n=236), 4,435 abstracts were screened and excluded, and 180 records were selected for full-text

screening. Of the 180 abstracts, a second screening for duplicates and non-retrievable texts removed 58 records, which resulted in the remaining 122 records being eligible for full-text screening.

Forty-five of 122 full texts were excluded due to the type of interventions, 27 were excluded due to study design (i.e. cross-sectional studies, or systematic reviews with no relevant included studies) and ten studies were ineligible because they did not use a control group. Additionally, eight full texts were not intervention studies, five studies contained the wrong study population, two were excluded as they took place outside of the school setting, and one study was not reported in English. In total, 98 full texts were excluded. Twenty-four studies met the inclusion criteria for this review, each of which represent a separate study. All 24 studies were included in the narrative review. Due to a lack of academic outcome data, only ten studies were included in the meta-analyses and four studies in the harvest plot. Brigman (2003) and Miller (2017) provided full outcome data for both types of analyses and so they were included in both data syntheses. Multiple attempts were made to secure missing effectiveness and equity data, but only one attempt was successful (Miranda, Brigman, & Peluso, 2007), which provided supplementary equity data for the Brigman (2003) study. Additionally, the review author searched for additional information, such as protocols, trial registration documents, online supplements, etc. for all 24 included studies. This search led to securing information for Eskreis-Winkler (2016) and Cohen (2006) via online supplementary materials but these documents do not serve as full reports themselves. The 2012 protocol for Tak (2014) was also located and used as supplementary material, but no other studies had a published online protocol. Thus, the search yielded a total of 25 full reports for 24 studies.

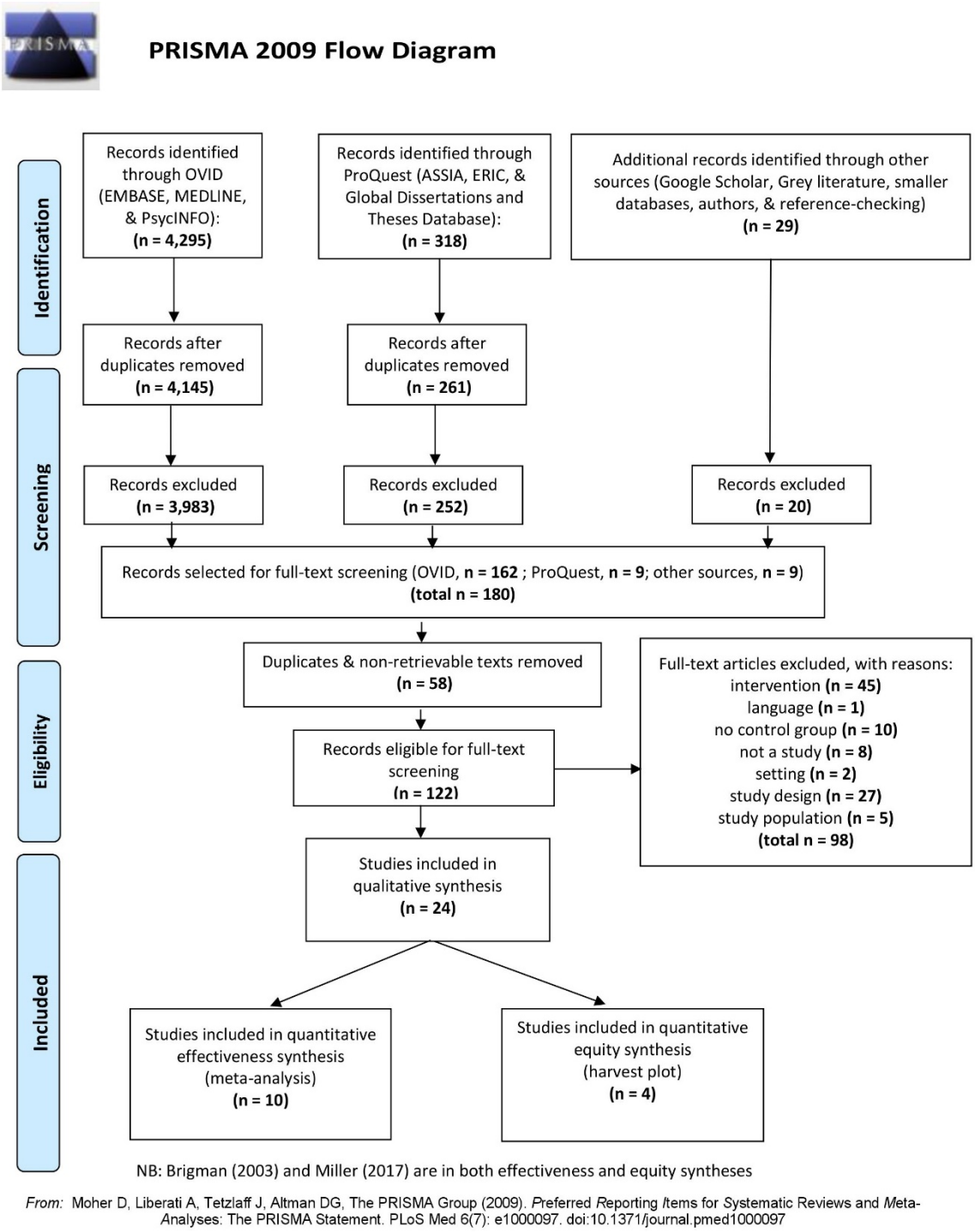


Figure 5. PRISMA Diagram.

4.1.2 Description of included studies

4.1.2.1 Study and sample characteristics

Of the 24 included studies, 10 (41.7%) were RCTs and 5 (20.8%) were clustered RCTs. Approximately half of the non-randomised studies used a matched group design (55%; n=5), one of which was a propensity score matching design (Dougherty, 2017) and the other four studies matched students based on important baseline demographic characteristics, such as sex, race, grade level. The remaining non-randomised studies used a pretest-posttest design (44%; n=4). The majority of the studies were conducted in the United States (n=18; 75%). Two studies were conducted in China (Harrison et al., 2017; Hu et al., 2011), one in Israel (Shoshani, Steinmetz, & Kanat-Maymon, 2016), Italy (Muratori et al., 2016), the Netherlands (Tak et al., 2014), and Pakistan (Arif, 2017). Across the studies, there were 21,009 participants in total, with the smallest study involving 30 students (DiGiacomo, 2014) and the largest study involving 10,880 students (Snyder, 2009). Interventions ranged in duration from a one-time lesson to 4 academic years, with the average, and most frequent, intervention duration being the medium-term length of 9 months (i.e. 37 weeks).

In term of sample characteristics, there were students reported at every grade through primary and secondary school, with the most common grade levels being in the middle school range (i.e. 6th through 8th grade or country-equivalent). Students ranged in age from 6 to 18 years, and the most frequently reported race was Caucasian (37.9%). Where reported in 21 of the 24 studies, gender was evenly split between male and female students when pooling these study samples. In the 14 studies that reported socioeconomic status, 52.6% of the students were considered 'low-income' (weighted average across

studies), though there was no consistent or clear method of measuring socioeconomic status.

4.1.2.2 Theories of change of included studies

Very few studies presented a programme theory in their reports. Due to the limited information available, the review author decided to synthesise the study findings by intervention content per protective factor in lieu of the original plan, which was to combine studies by similar theories of change under each protective factor as noted in the methods section.

4.1.2.3 Intervention programme details

Protective factors. Grouping interventions by the predominant psychosocial protective factor, most studies focused on school-based social support as the protective factor of interest (n=9; 37.5%), followed up by academic self-efficacy and self-regulation, which includes self-regulated thinking (n=6; 25%). Three interventions were based on improving cognitive style (12.5%), and two interventions each used academic motivation and coping skills (8.3% each). Academic planning and academic perseverance were targeted in one study each (4.17% each), and no studies directly utilised adaptive help-seeking or self-care behaviours in their interventions. Virtually all comparison groups received treatment as usual, which was a normal school curriculum, so these data were not included in the table of study characteristics in table 7 below.

Programme content. There were a host of intervention programme components and processes used in the body of evidence. Half of the included studies noted the use of curriculums or manualised interventions. Brigman (2003) used the Student Success Skills programme, which is a classroom guidance and counselling intervention. Dougherty (2017) used the Reconnecting Youth programme, an intervention that increases academic

achievement and reduces substance abuse. DuPuis (2013) and Portwood (2005) both evaluated the YouthFriends programme for school-based mentoring. The ChildCARE intervention in Harrison (2017) is a community-level intervention to improve social support and coping strategies for high-risk students. Hu (2011) developed the 'Learn to Think' (LTT) curriculum to adjust students' self-regulated thinking skills. Miller (2017) assessed the effectiveness of the Relationship Building Intervention for social support among 5th graders. Muratori (2016) used the coping skills intervention, Coping Power. The Maytiv Programme used in Shoshani (2016) aims to improve academic motivation and social support. Snyder (2009) used the Positive Action intervention, which is a school-wide intervention that targets self-concept. Tak (2014) used the Op Volle Kracht programme, a resiliency training programme to improve self-regulation, and Terry (2016) used the Footprints intervention for academic motivation.

Arif (2017), DiGiacomo (2014), Eggert (1990), and Paunesku (2015) evaluated the effect of several school-based activities such as exercises pertaining to specific protective factors and group discussions. Six interventions focused more on changing learning contexts and functions instead of using specific intervention activities (Asmus, 2017; Curran, 2018; Felner, 1982; Haynes, 1990; Karcher, 2008; and Reyes, 1994). For example, Curran (2018) grouped students by mathematic ability to evaluate if changing the context in which they receive maths instruction could improve their academic self-concept. Both Cohen (2006) and Eskreis-Winkler (2016) used brief intervention tasks, such as computerised or handwritten prompts.

Implementation strategies. The following 9 studies utilised an interactive classroom delivery method which encouraged participation between the students and the teacher: Arif (2017), Brigman (2003), Curran (2018), Dougherty (2017), Felner (1982), Miller (2017), Muratori (2016), and Shoshani (2016). Of the remaining 15 studies, Eggert

(1990) and Harrison (2017) implemented their programmes in small peer groups. Eleven studies delivered the intervention individually (Asmus, 2017; Cohen, 2006; DiGiacomo, 2014; DuPuis, 2013; Eskreis-Winkler, 2016; Hu, 2011; Karcher, 2008; Paunesku, 2015; Portwood, 2005; Reyes,1994; and Terry, 2016). Both Haynes (1900) and Snyder (2009) used a school-wide implementation strategy.

Prevention level. The body of evidence presented multiple prevention types across the academic risk continuum. The 15 following studies were universal programmes: Arif (2017), Cohen (2006), Curran (2018), DiGiacomo (2014), DuPuis (2013), Eskreis-Winkler (2016), Haynes (1990), Hu (2011), Miller (2017), Muratori (2016), Paunesku (2015), Portwod (2005), Shoshani (2016), Snyder (2009), and Tak (2014). Six programmes were targeted interventions (Asmus, 2017; Brigman, 2003; Dougherty, 2017; Karcher, 2008; Reyes, 1994; and Terry, 2016). Finally, Eggert (1990), Felner (1982), and Harrison (2017) were indicated prevention programmes.

4.1.2.4 Effectiveness and equity outcome measures

Twenty two of the 24 studies reported some form of academic data. However, the effectiveness outcome data from 12 studies could not be included in a standardised mean difference meta-analysis because they either omitted simple statistics, like standard deviations, or reported information that cannot be included in this type of forest plot, such as beta coefficients. Eighteen studies were not applicable in the equity data analyses because they did not report within-study moderator effects or interaction effects.

Of the 10 studies that reported usable academic outcome data for meta-analyses, four (40%) studies reported grade point averages as the outcome measure of interest (Miller, 2017; Muratori, 2016; Shoshani, 2016; Tak, 2014). An additional four (40%) studies reported standardised test scores (Brigman, 2003; Curran, 2018; Hu, 2011; Snyder, 2009).

One (10%) study reported individualised maths test scores (DiGiacomo, 2014) and one study (10%) reported course-specified marks in Chinese and maths (Harrison et al., 2017). Across the four studies used in harvest plots, Brigman (2003) reported standardised math exam scores by race. Eskreis-Winkler (2016) reported math scores from an independent, online educational curriculum, Khan Academy, by gender and race. Cohen (2006) reported outcome data on GPAs by race and gender. Miller (2017) reported outcome data on GPAs as well by gender and race.

4.1.2.5 Attempts to secure incomplete or missing outcome data

The authors of all 14 studies that had missing or incomplete effectiveness data (e.g. means, standard deviations, etc.) for academic performance outcomes were contacted. Additionally, the authors of the 20 studies that had missing or incomplete academic performance equity data were emailed as well. Authors from approximately half (n=11) of the studies responded. One response (Brigman, 2003) was successful, and the review author was able to secure vital equity data. The other ten responses proved unsuccessful. Arif (2017) stated that they never recorded academic outcome data as the school administrators in Lahore refused access to students' school records. Asmus (2017) had no additional information for this study, so it was excluded from both the equity and effectiveness data syntheses. Both Curran (2018) and Snyder (2009) did not measure equity factors at the student level, and thus there was no subgroup data on academic performance by PROGRESS-Plus factors. Eskreis-Winkler (2016) and Karcher (2008) both did not have the time to locate the requested data. Haynes (1990), Portwood (2005), and Reyes (1994) no longer had access to the data from their respective studies. Shoshani (2016) did not perform moderator analyses. Terry (2016) responded to the initial request but has yet to provide any data. The authors from four studies never responded (Dougherty, 2017;

DuPuis, 2013; Eggert, 1990; Paunesku 2015) and contact information for the main author of one study could not be found (Felner, 1982).

Study ID	Study Details	Sample size	Participants	Intervention description	Protective factor(s)	Academic Outcome Measure & Data	PROGRESS-Plus Data	Study Quality
30. Arif 2017	RCT, Pakistan, duration n.s., Universal	64 students	9th and 10th graders in Lahore; age: 14-16y; gender: 100% male; SES: n.s.	classroom activities fostering creativity, self-esteem, self-efficacy, internal locus of control, autonomy, problem-solving skill, sense of optimism and hope, stress coping skills and teacher-student relationship.	Self-efficacy and self-regulation Coping Social Support Perseverance Cognitive style			Low risk (RoB 2.0)
32. Asmus 2017	RCT, USA, 15 weeks, Targeted	95 students	Cognitively disabled secondary students; age: n.s.; gender: 68.8% male; race: 70% white, 15% black, 8% Hispanic, 4% Asian, 0.2% Native American, and 4% Mixed race; SES: 31% low-income	a group of three to six peers formed a network for each focus student. Peers exchanged social interactions and support with their fellow student with disabilities	Social Support			Some concerns (RoB 2.0)
86. Brigman 2003	Matched group design, USA, 9 months, Targeted	360 students (30 students across 12 schools)	Florida elementary, middle, and high schools; age: n.s.; gender: 54% female; race: 67.6% White, 22.3% Black, and 10% Latino; SES: n.s.	The student success skills curriculum for group counselling and classroom guidance focused on cognitive, social, and self-management skills. The group counselling intervention consisted of 8 weekly sessions of 45 minutes each. Goal-oriented tasks to improve academic achievement were tracked using a chart to see if students applied skills they learned.	Self-efficacy and self-regulation Academic Planning	FCAT MATH 2002 (standardised test score) intervention (M=662.46, SD=44.797, N=97); Control (M=656.67, SD=28.696, N=125)	Race (MANCOVA)	Critical risk (ROBINS-I)

Study ID	Study Details	Sample size	Participants	Intervention description	Protective factor(s)	Academic Outcome Measure & Data	PROGRESS-Plus Data	Study Quality
19. Cohen 2006	RCT, USA, one-time intervention, Universal	133 students	7th grade students; age: n.s.; 55% female; race: 38% black, 46% white, 11% Hispanic, and 5% Asian American; SES: low- and middle-income	Students chose a protective factor that they identify the most with. They were then instructed to write about why this value was most important to them, and also to describe why someone else would view this value as beneficial as well	Cognitive style	GPA n/a	Race & Gender (linear regression model)	High risk (RoB 2.0)
22. Curran 2018	Pretest posttest with control, USA, 8 months, Universal	282 students	4th and 5th grade students; age: 9-11 yrs.; gender: n.s.; race: 60% White, 19% Asian, 11% Hispanic, 6% Black, and 5% multiracial; SES: 19% low-income	The intervention was designed to improve students' beliefs and flexibility about their math skills. Students were assessed pre-intervention and then sorted by ability into assigned math classrooms. Students received daily math lessons in these groups and then returned to their homeroom for the remainder of the day.	Cognitive style	STAR math score (standardised test score) intervention (M=762, SD=82, N=140); Control (M=762, SD=66, N=142)		Critical risk (ROBINS-I)
73. DiGiacomo 2014	RCT, USA, 3 weeks, Universal	30 students	6th and 7th grade students; age: n.s.; 50% male; race: n.s.; SES: upper middle class	1) Students were trained in various self-regulated learning strategies (e.g., planning, monitoring, reflection) 2) Math review questions. While solving these questions they were asked to make predictive and postdictive confidence judgments focusing on	Self-efficacy and self-regulation (self-regulated learning)	math performance (individual test score) intervention (M=4.6, SD=1.99, N=15); Control (M=2.8,		Low risk (RoB 2.0)

Study ID	Study Details	Sample size	Participants	Intervention description	Protective factor(s)	Academic Outcome Measure & Data	PROGRESS-Plus Data	Study Quality
				how well they thought they would be able to/were able to complete each question. 3) Students charted discrepancy between their judgments and their actual performance to track calibration accuracy. 4) Reflection worksheets on how they could improve their weaknesses		SD=1.44, N=15)		
76. Dougherty 2017	Propensity score matching design, USA, 9 months, targeted	110 students	middle school and high school students; age: 12-17 yrs.; gender: 36% female; race: 76% Latino, 18% White, 4% American Indian, and 2% Black; SES: n.s.	Reconnecting Youth programme consists of a class with goals of improving academic achievement, reducing substance use, and improving mood management. RY classes contain 10–12 students. The curriculum delivers 79 lessons. The initial unit introduces students to the RY model; the next four units are on self-esteem, decision-making, personal control, and interpersonal decision-making.	Social Support	GPA n/a		Low risk (ROBINS-I)
9. DuPuis 2013	Pretest posttest with control, USA, 9 months, Universal	83 students	middle school students; age: 10.9-14.9 yrs.; 53% males; race: 15% Black; 72% White; and 13% Hispanic; SES: n.s.	YouthFriends is a school-based mentorship programme. Students were placed in the treatment group based on their level of need or if they already had a mentor. Students who did not already have a mentor or were referred to the programme	Social Support	GPA n/a		Low risk (ROBINS-I)

Study ID	Study Details	Sample size	Participants	Intervention description	Protective factor(s)	Academic Outcome Measure & Data	PROGRESS-Plus Data	Study Quality
				shortly after the mentoring pairs were established were used as a waitlist control group				
99. Eggert 1990	Matched group design, USA, 4.5 months, Indicated	264 students	secondary students; age: 14-19 yrs. (M = 16.6 yrs.); gender: 43% female; race: mostly white; SES: middle class	Four days/week: group discussion of psychosocial problems and skills training in interpersonal communication, problem solving, decision making, and self-management (40 min); supervised study/peer tutoring (15 min). One day/week: reviewing students' progress in classes, goal setting, planning drug-free, weekend activities. Two half days/month: visiting community agencies for GED programmes, community colleges, employment security offices, and recreational activities of the students' choice	Social Support	GPA n/a		Serious risk (ROBINS-I)
25. Eskreis-Winkler 2016	RCT, USA, one-time, 25 mins intervention, Universal	209 students	5th and 6th graders; age: n.s.; gender: 44% female; race: 53% White, 23.9% Black, 11% Hispanic, 9.1% Asian, and 5.3% mixed race; SES: n.s.	The treatment condition consisted of two parts. In Part I, students learned the tenets of deliberate practice: 1) Focus on weaknesses, 2) Get feedback, 3) Concentrate 100%, and 4) Repeat until mastery. Student were given activity prompts, an illustrative video, and a letter-writing exercise. Part II was motivational. It targeted expectancies and values,	Academic motivation	Khan academy performance (individualised test score) n/a	Gender and race (linear regression model)	Low risk (RoB 2.0)

Study ID	Study Details	Sample size	Participants	Intervention description	Protective factor(s)	Academic Outcome Measure & Data	PROGRESS-Plus Data	Study Quality
				(expectancy-value theory). To address expectancies, the module taught the importance of investing effort in deliberate practice. To address costs, a component of value in EVT, the module encouraged students to interpret feelings of being challenges as indicators of engaging with current weaknesses. Thus, the intervention targeted practice-specific expectancies and costs in an attempt to motivate deliberate practice.				
38. Felner 1982	Matched group design, USA, academic year, Indicated	185 students	high school students; age: n.s.; gender: n.s.; race: 57% black, 19% white, 22% Hispanic, 2% other; SES: low-income	All treatment freshmen were assigned to one of four Project Home rooms. The teachers oversaw most of the intervention, helping students pick classes, counselling them through difficulties, etc. The objective for this component was three-fold: 1. Increase structural social support from a school-based agent, i.e. the teacher. 2. Increase student's feeling of accountability in their school environment. 3. Provide students with school information, regulations, and expectations. All Project students were assigned to	Social Support Cognitive style	academic adjustment measures (GPA) n/a		Low risk (ROBINS-I)

Study ID	Study Details	Sample size	Participants	Intervention description	Protective factor(s)	Academic Outcome Measure & Data	PROGRESS-Plus Data	Study Quality
				the same core classes to create peer support and sense of belonging.				
8. Harrison 2017	cRCT, China, 4 years, Indicated	790 students (395 in child-only and condition arms)	primary and secondary students; age: 6-17 yrs.; gender: 51.6% male; race: 100% Asian; SES: 100% low-income	The ChildCARE intervention consists of programming at three levels: Child, caregiver, and community. Small peer-group activities foster resilience characteristics for children affected by parental HIV/AIDS, including positive thinking, emotional regulation, coping skills, problem solving, support seeking, positive future orientation, and enhanced self-esteem. The intervention was delivered to peer groups based on child's grade level.	Coping Social Support Cognitive style	academic performance (course-specific marks) child-only intervention (M=3, SD=1.329, N=200); control (M=2.891, SD=1.309, N=195)		Low risk (RoB 2.0)
66. Haynes 1990	Pretest posttest with control, USA, 9 months, Universal	174 students	4th and 6th grade students: age: n.s.; gender: 51% female; race: 100% Black; SES: 100% low-income	The development of comprehensive school plan which includes academic, social, and staff development goals, a Mental Health Team to address schoolwide issues and individual student and staff concerns, and a parent programme to include families in different levels of school life.	Self-efficacy and self-regulation (incl. self-regulated learning) Social Support Cognitive style			No information (ROBINS-I)
29. Hu 2011	RCT, China, 4 years, Universal	166 students	1st - 3rd grade students; age: M=6 yrs.; gender: 54.2%	Manualised intervention including 16 activities covering different course content and methods, with	self-regulated thinking	academic achievement test		Low risk (RoB 2.0)

Study ID	Study Details	Sample size	Participants	Intervention description	Protective factor(s)	Academic Outcome Measure & Data	PROGRESS-Plus Data	Study Quality
			male; race: 100% Chinese; SES: n.s.	thinking qualities trained in each activity. For example: thinking methods include step-by-step observation, imagination, and story-inventing in a literature course		(standardised test score) intervention (M=93.28, SD=3.83, N=90); Control (M=86.22, SD=8.09, N=76)		
118. Karcher 2008	RCT, USA, 8 weeks, Targeted	525 students	primary and secondary school students; age: 10-18 yrs.; gender: 66% girls; race: 49% Hispanic, 1% white, 8% Black, 0.3% Asian, 30% Hispanic/Anglo biracial; SES: low-income	Intervention students were matched with a youth mentor for weekly school-based mentorship	Social support	GPA n/a		Some concerns (RoB 2.0)
62. Miller 2017	Pretest posttest with control, USA, 26 weeks, Universal	627 students (29 classrooms)	5th grade students; age: M=10.08 yrs.; gender: 52.5% female; race: 55.5% European American, 9.4% Hispanic/Latino, 8.0% Asian American, 3.8% Black/Black, 1.1%	The RBI consists of 21 activities across five units: Diversity and Inclusion, Empathy and Critical Thinking, Communication, Problem Solving, and Peer Relationships. Activities promote active involvement by incorporating games, hands-on activities, role plays, and group discussions. All activities occur within	Social support	academic achievement (GPA) intervention (M=3.64, SD=0.05, N=243); Control (M=86.22,	Gender and race (linear regression model)	Low risk (ROBINS-I)

Study ID	Study Details	Sample size	Participants	Intervention description	Protective factor(s)	Academic Outcome Measure & Data	PROGRESS-Plus Data	Study Quality
			Native American, 0.2% Pacific Islander, 17.9% Multiracial; SES: upper middle class	the peer context; students learn, practice, and reflect on concepts in dyads, small groups, and within the whole classroom.		SD=0.05, N=359)		
117. Muratori 2016	cRCT, Italy, 6 months, Universal	184 students (9 classrooms)	1st and 2nd grades: age: M=7.5 yrs.; gender: 48% male; race: 82% Italian, other races n.s.; SES: n.s.	Coping Power is a classroom intervention curriculum which consists of 24 session, some of which include lessons on behavioural goal setting, self-control, coping, perspective taking and problem solving etc.	Coping	GPA intervention (M=8.29, SD=0.94, N=100); Control (M=7.93, SD=0.7, N=62)		Low risk (RoB 2.0)
51. Paunesku 2015	RCT, USA, two weeks, Universal	1,594 students	High school students; age: n.s.; gender: 50% male; race: 33% Hispanic, 23% white, 17% Asian, 11% Black, 15% mixed race; SES: approx. 50% low-income	Students read an article describing the brain's ability to grow and reorganise itself as a result of using strategies on challenging tasks, which is meant to promote academic motivation and shift their cognitive style. This message is then reinforced through two writing exercises. Secondly, the sense-of-purpose task is to help students articulate how schoolwork could help them accomplish meaningful, beyond-the-self life goals. In the intervention, the students were instructed to write about how they wished the world could be a better place and then told to think about	Academic motivation Cognitive Style	GPA n/a		Low risk (RoB 2.0)

Study ID	Study Details	Sample size	Participants	Intervention description	Protective factor(s)	Academic Outcome Measure & Data	PROGRESS-Plus Data	Study Quality
				their own goals and how in school performance could help them achieve these goals				
48. Portwood 2005	Matched group design, USA, 9 months, Universal	208 students	4th through 12th grade students; age: n.s.; gender: 52% female; race: full data not available; SES: n.s.	YouthFriends activities took the form of weekly 1-hour mentoring sessions throughout the academic year	Social support	GPA n/a		Serious risk (ROBINS-I)
100. Reyes 1994	RCT, USA, 1 year, Targeted	145 students	8th grade students; age: M=14.4 yrs.; gender: 57% female; race: 77% Hispanic, other races n.s.; SES: 95% low-income	Peer Support Component of the intervention eases students' transition to high school through high school-based peer support. Eighth graders were computer-matched with same-gender Community High School student 'peer helpers,' based on academic, social, recreational, and extracurricular interests. 8th grade students participated in high school-based activities with their peer helpers throughout the academic year	Social support	academic adjustment measures (GPA) n/a		Low risk (RoB 2.0)
42. Shoshani 2016	cRCT, Israel, 9 months, Universal	2,517 students (70 classrooms)	7th through 9th grade students; age: 11.9-14.9 years (M=13.54 yrs.); race: 98% Israeli; gender: 50% female; SES: 66%	Maytiv Programme consists of 5 components. Step 1: Stories (20–30 min). Students are exposed to 2 to 3 stories in each lesson and learn to derive certain lessons and action items from them.	Academic motivation Social support Perseverance	school achievement (GPA) intervention (M=81.15, SD=9.35, N=1110);		Low risk (RoB 2.0)

Study ID	Study Details	Sample size	Participants	Intervention description	Protective factor(s)	Academic Outcome Measure & Data	PROGRESS-Plus Data	Study Quality
			middle SES, 16% high SES, and 18% low SES.	Step 2: Exercises (30–40 min). In each lesson, the students are given 1 to 2 exercises to be completed in pairs, small groups, or as a class. Based on the notion of learning through imitation, many of the exercises involve acting They are also exposed to other activities ranging from guided visualisation to team-building exercises, cognitive behavioural therapy techniques, sharing personal story, and drawing. Step 3: Group discussion of previous activities. Step 4: writing prompt Step 5: Putting to use what they've learned through a school-based action of their choosing		Control (M=77.69, SD=12.32, N=1098)		
14. Snyder 2009	cRCT, USA, 4 years, Universal	10,880 students (approx. 544 students from 20 schools)	Kindergarten-6th grade students; age: n.s.; gender: n.s.; race: 1.79% Black, 2.05% Chinese, 11.61% Filipino, 5.61% Hawai'ian, 2.45% Hispanic, 2.02% Indochinese, 4.26% Japanese, 1.19% Korean,	The Positive Action curriculum consists of 140 lessons per grade, per academic year, offered in 15 to 20 min by classroom teachers. Lessons cover six major units on topics related to self-concept, physical and intellectual actions, self-control, getting along with others, being honest with yourself and others, and continuous self-improvement. The classroom curricula utilise an interactive	Cognitive style	HCPS II (Hawaii state standardised exam) intervention (M=41.89, SD=15.59, N=9 schools); Control (M=26.67, SD=7.79, N=9 schools)		Low risk (RoB 2.0)

Study ID	Study Details	Sample size	Participants	Intervention description	Protective factor(s)	Academic Outcome Measure & Data	PROGRESS-Plus Data	Study Quality
			0.44% Native American, 31.86% Part Hawai'ian, 1.41% Portuguese, 3.11% Samoan, 17.52% White, 14.6% Other; SES: 54.32% low-income	approach, whereby interaction between teacher and student is encouraged through the use of structured discussions and activities, and interaction between students is encouraged through structured or semi-structured small-group activities, including games, role plays, and practice of skills				
58. Tak 2014	cRCT, the Netherlands, 5 months, Universal	1,341 students (9 schools)	8th grade students; age: M=13.9 yrs.; gender: 47.3% female; race: 83.1% Dutch; 16.9% other; SES: n.s.	Op Valle Kracht intervention consists of 16 lessons of 50 minutes. The first 8 lessons explain/utilise the CBT-derived principles. Lessons 9 to 16 are directed at social and coping skills, self-esteem, problem solving, and decision-making. During every lesson, the theory behind the techniques is shortly explained followed by actively practicing the skills by students. Several means are used for teaching, such as discussions, role-plays, and skits. Group leaders are responsible for group atmosphere and cohesion, and they will guide discussions and role-plays. Each lesson includes homework for the next lesson. Practicing skills facilitates internalisation of these	Self-efficacy and self-regulation (incl. self-regulated learning)	GPA intervention (M=7.07, SD=1.7, N=634); Control (M=7.22, SD=1.4, N=707)		Some concerns (RoB 2.0)

Study ID	Study Details	Sample size	Participants	Intervention description	Protective factor(s)	Academic Outcome Measure & Data	PROGRESS-Plus Data	Study Quality
				skills, which results in better programme effects				
20. Terry 2016	RCT, USA, 8 weeks, Targeted	43 students	middle school students; age: n.s.; gender: 61% male; race: 52.3% Black, 38.1% White, and 9.6% Asian/Pacific Islander or Hispanic; SES: 49.2% low-income	Footprints consists of psychoeducation, values clarification, individual goal-setting, behavioural activation, cognitive restructuring, calming/problem solving, tailored individual support, and developing plans to increase protective factors	Academic motivation	GPA n/a		Low risk (RoB 2.0)

Table 7. Characteristics of included studies. N.S. = not specified. N.A. = not applicable; as in some studies reported either incomplete outcome data or data not appropriate for meta-analyses (i.e. correlations, beta coefficients, etc).

4.1.3 Excluded studies

The main reasons for exclusion of studies can be found PRISMA design under section 4.1.1. Given the number of excluded studies (n=98), the characteristic of excluded studies table has been added to the appendix (see appendix 7.1).

4.2 Risk of bias in included studies

Risk of bias in included studies has been separated by quality assessment method in this review. For both individual and clustered randomised controlled trials, risk of bias is summarised using the Cochrane Risk of Bias tool (RoB 2.0). For non-randomised studies, the risk of bias for each study is synthesised below using the Cochrane Risk of Bias of Included Non-Randomised Studies (ROBINS-I). Risk of bias graphs and summary tables for randomised and non-randomised studies can be found below section 4.2.1 (figures 6 and 7) and section 4.2.2 (figures 8 and 9), respectively. As noted in the methods section, the serious and critical risk of bias for the ROBINS-I are symbolised by the red minus sign in the figures below.

4.2.1 For randomised controlled trials using the RoB 2.0

4.2.1.1 Random sequence generation

Low risk. Both Eskreis-Winkler (2016) and Paunesku (2015) were rated as low risk because students were randomly assigned using an online algorithm to deliver the interventions. Terry (2016) used a list of randomly-generated numbers to assign students to treatment conditions.

Unclear risk. Of the 15 RCTs, 10 studies were rated as unclear risk for random sequence generation: Asmus (2017), Arif (2017), DiGiacomo (2014), Harrison (2017), Hu

(2011), Karcher (2008), Muratori (2016), Reyes (1994), Shoshani (2016), and Snyder (2009). They all reported that they conducted an RCT but did not specify their randomisation methods.

High risk. Both Cohen (2006) and Tak (2014) had baseline imbalances that suggested issues with their randomisation processes, as well as no information about their random sequence generations, so these studies was considered high risk.

4.2.1.2 Allocation concealment

Low risk. Like their randomisation processes, the treatment conditions were concealed through the respective online platforms, so Eskreis-Winkler (2016) and Paunesku (2015) were found to be low risk. Most studies did not specify their allocation concealment methods but there was no evidence to suggest that participants were aware of their treatment condition until after they were recruited and assigned. The rating of low risk of bias applied to the following studies: Arif (2017), Asmus (2017), Cohen (2006), DiGiacomo (2014), Harrison (2017), Hu (2011), Karcher (2008), Muratori (2016), Reyes (1994), Shoshani (2016), Tak (2014), and Terry (2016).

Unclear risk. Snyder (2009) reported that randomisation of schools occurred before recruitment, so it is unknown if students were recruited into the study with the prior knowledge of their treatment allocation.

4.2.1.3 Blinding of participants and personnel

Low risk. In most cases, participants and delivery staff cannot be completely blinded, and may have some awareness of a study taking place. However, they may be unaware of deliberate efforts to test programme effectiveness through different treatment conditions, which reduces the likelihood of bias. By this logic, in Arif (2017) and Asmus

(2017), participants and/or personnel were assumed to be aware of intervention allocation but they were considered low risk because their possible knowledge could not have biased academic performance outcomes since both studies had no access to academic data.

Cohen (2006) clearly noted that their study was a double-blind RCT whereby students were randomly assigned at the beginning of the school year, and both teachers and students remained unaware of treatment assignment throughout the intervention. Similarly, Eskreis-Winkler (2016) and Paunesku (2015) were also double-blind computer interventions. The respective participants in both the Hu (2011) and DiGiacomo (2014) studies were assumed to be unaware of their treatment conditions but the researchers, who were involved with programme implementation, were likely aware of treatment allocation. These studies were given a low risk of bias as the researchers' presumed knowledge did not appear to influence how the interventions were intended to be delivered. The teachers who delivered the intervention in the Muratori (2016) study were not blinded to intervention assignment but they were not informed that grades were recorded as outcome data, so they probably did not bias the outcome of interest.

Unclear risk. Harrison (2017)'s students were assumed to be blinded; however, the trial personnel were involved with implementing the programme for all condition arms, so this study was deemed as having an unclear risk of bias. In Snyder (2009) and Tak (2014), the teachers in the respective intervention arms were fully aware of students' intervention assignment. Students were likely unaware of their treatment allocation and there was no risk of contamination as both studies were school-level cRCTs.

High risk. Karcher (2008) was given a high risk of bias because neither participants nor personnel were not blinded to the intervention. Students receiving the mentoring intervention were aware of their treatment assignment because they were either

referred or self-nominated for entry into the programme. It is also assumed that the trial personnel were aware of treatment assignment because the mentors had to meet with students on a regular basis. Similarly, the intervention students and personnel in Reyes (1994) were involved with a series of activities that would make it apparent that they were allocated to the treatment arm of a mentoring programme. The delivery staff in both Shoshani (2016) and Terry (2016) were not blinded to the intervention. Students were also informed that they would be randomly assigned to either treatment or control conditions.

4.2.1.4 Blinding of outcome assessors

Low risk. For studies that had no access to academic data, it would not be possible to introduce bias, so a low risk of bias rating was given to Arif (2017) and Asmus (2017). Eskreis-Winkler (2016) was considered low risk as all personnel were blinded to the computer-based intervention. Additionally, several studies were deemed low risk because the outcome assessor was blinded to the intervention (Cohen, 2006; Harrison, 2016; Karcher, 2008; Muratori, 2016; Paunesku, 2015; Reyes, 1994; Snyder, 2009; Tak, 2014). For example, Karcher (2008) was assumed to have a low risk of bias as data on school grades were reported by a third-party case manager. Additionally, Snyder (2009) study was determined to be low risk as well because academic outcome data were collected by the Hawai'ian Department of Education independently of the study, so it is unlikely that these administrative data could have tampered with.

DiGiacomo (2014), Hu (2011), Shoshani (2016), and Terry (2016) were not blinded as the outcome assessors, but it was determined that knowledge of treatment allocation probably would not result in bias, and as such, they were all considered low risk.

4.2.1.5 Incomplete outcome data

Low risk. Studies with an attrition rate of 0-10% were considered as low risk as stated by the RoB 2.0 (Higgins et al., 2016). This rationale applied to the following studies: Cohen, 2006; DiGiacomo, 2014; Eskreis-Winkler, 2016; Harrison, 2017; Hu, 2011; Karcher, 2008, and Muratori, 2016). Studies that had more than 10% attrition but used analytical methods to account missing outcome data were also rated as low risk (Paunesku, 2015; Shoshani, 2016; Tak, 2014; Terry, 2016). For example, Tak (2014) averaged a moderately high attrition rate of approximately 15% across all five post-intervention follow up timepoints. However, they used multiple imputation to account for missing data.

Arif (2017) and Asmus (2017) were rated as low risk as attrition could not have affected outcomes due to their inability to secure academic data. Snyder (2009) could not obtain academic outcome data from the Hawai'ian Department of Education for 50% of their school sample (n =20 schools). However, it was rated as low risk because the proportions and reasons for missing data were the same for both intervention and control schools due to the matched group design.

High risk. Reyes (1994) was considered as high risk because they noted 'significant loss of data over time [which] warranted consideration of how representative the final sample at follow-up 2 was of the original sample at baseline' (p. 358).

4.2.1.6 Selective reporting

Both effectiveness and equity data reporting were weighted equally in determining risk of bias for this section. All studies except Tak (2014) did not have any published protocols with which reporting on outcomes could be compared to identify potential selective reporting. Therefore, decisions on selective reporting were left to the review author's judgements.

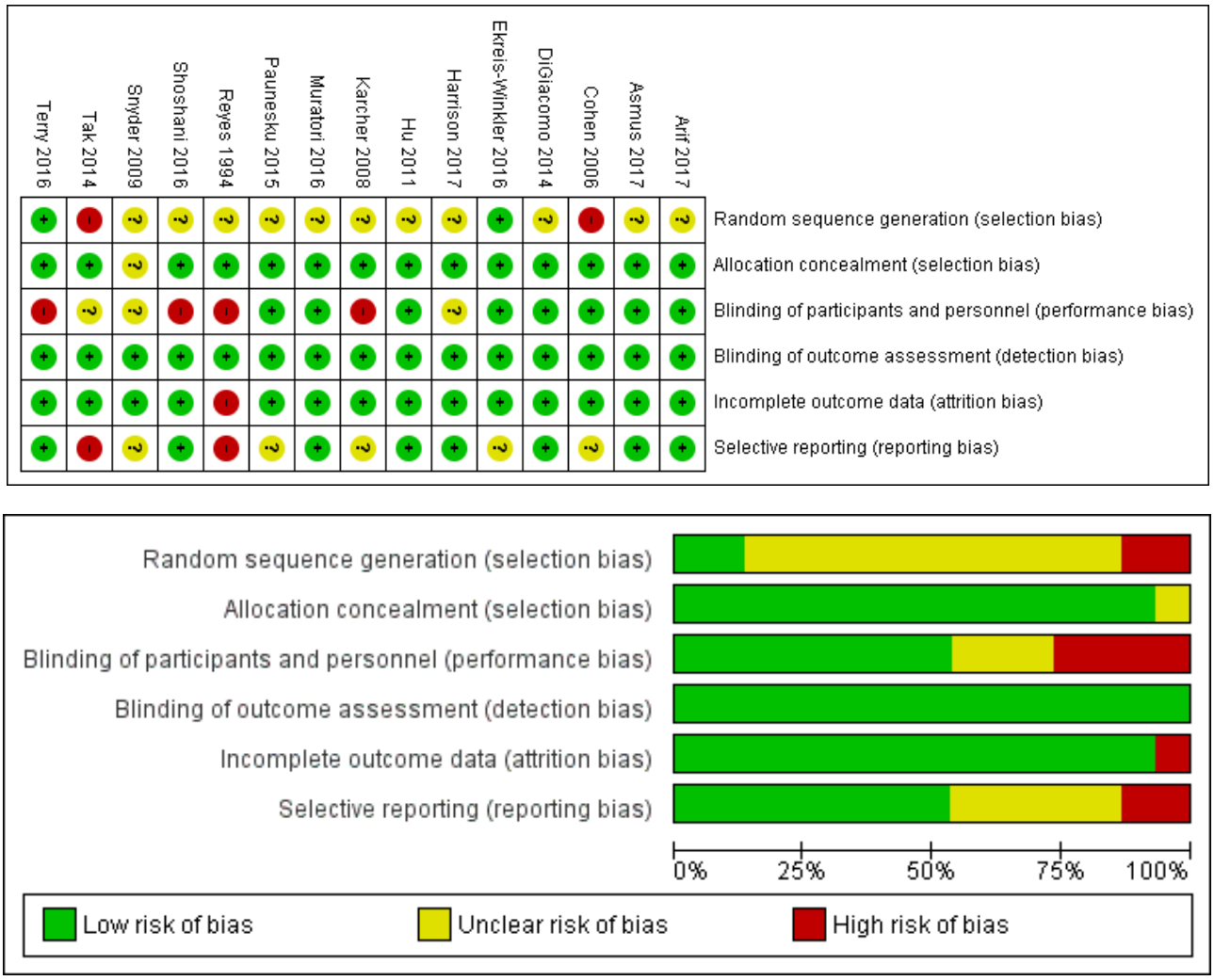
Low risk. Arif (2017) and Asmus (2017) were given a low risk of bias rating on selective reporting as these studies had no academic achievement data to report for either effectiveness or equity outcomes. Arif (2017) stated via email that they were refused access to school grades and Asmus (2017) noted that they never recorded academic outcome data. Muratori (2016), DiGiacomo (2014), and Terry (2016) were all given a low risk of bias for selective reporting as they fully reported effectiveness data. There was no reason to suggest that any of these studies intentionally excluded equity findings due to negative results. Both Hu (2011) and Harrison (2017) provided enough relevant effectiveness to be considered low risk of bias, and like the previously mentioned studies, they were believed to not have purposely excluded equity data. Shoshani (2016) reported full effectiveness outcome data and noted in their correspondence that they have no additional equity data to report, so they were given a low risk of bias rating.

Unclear risk. All studies in this section were rated as unclear risk as they were missing either effectiveness or equity data without clear reasoning as to why this evidence was not reported. Cohen (2006) and Eskreis-Winkler (2016) both report outcome data on equity effects by race and gender, respectively, using beta coefficients. However, they chose not to report basic information like pre- and post-intervention means and standard deviations by treatment condition, which are essential to calculating effect sizes by standardised mean difference.

Karcher (2008) and Paunesku (2015) received an unclear risk of bias as they were missing complete, basic data like means and standard deviations for effectiveness data syntheses but instead report beta coefficients and odd values, respectively. Additionally, both authors reported detailed equity-relevant data, such as demographic information by race, gender, and socioeconomic status, but it remains unclear whether these equity-

relevant data were used in equity analyses that would have produced equity evidence suitable for this review. Similarly, Snyder (2009) was considered as unclear risk as they report thorough equity-relevant data on demographics, which could suggest the possibility of equity data analyses. Without the requested data from Snyder (2009) or the other authors that received an unclear risk rating, the review author is unable to determine whether or not equity and effectiveness evidence were purposefully excluded from their findings.

High risk. Reyes (1994) was rated as high risk because they reported the post-intervention means and standard deviations on GPA but do not specify if these data refer to the treatment or control group, or both combined. They posit that students experienced significant GPA improvements between baseline and post-intervention, but there is no clear way of assessing this claim given the selective reporting. Tak (2014) was considered as high risk in terms of equity outcome data. As noted in their protocol (Tak et al., 2012), the researchers intended to account for demographic variables such as gender and ethnicity in their regressions across their outcomes of interest. In the 2014 full report, the model is missing data for academic outcomes by race and gender. It appears that the researchers may have chosen to exclude these data from the analysis presented. Figures 6 and 7 below present the risk of bias summary and graph for the randomised trials, respectively.



Figures 6 & 7. Risk of bias summary and graph (RoB 2.0).

4.2.2 For non-randomised studies using the ROBINS-I

4.2.2.1 Confounding

Low risk. Brigman (2003), Dougherty (2017), Eggert (1990), Felner (1982), and Portwood (2005) were all rated as low risk as they utilised a matched group design to minimise selection bias across important baseline variables. Even though they used a non-equivalent comparison group in their pretest-posttest design, the significance testing in DuPuis (2013) indicated that there were minimal baseline differences between the intervention and control groups.

Both Haynes (1990) and Miller (2017) also used a pretest-posttest design. Haynes (1990) controlled for gender, socioeconomic status, race, achievement level, and school level at baseline as well as in their MANCOVAs, so it was rated as a low risk study. Miller (2017)'s chi-square analyses revealed significant and substantive differences between the treatment and control conditions for the variables of household income, parents' education, and student ethnicity, suggesting issues with selection bias. However, the data were adjusted for using those variables in a multilevel model regression, so this study was considered to be low risk.

Critical risk. Curran (2018) does not control for confounding variables at baseline or in the data analysis, so there is a critical risk of selection bias.

4.2.2.2 Classification of intervention (treatment differentiation)

Low risk. All nine studies (Brigman, 2003; Curran, 2018; Dougherty, 2017; DuPuis, 2013; Eggert, 1990; Felner, 1982; Haynes, 1990; Miller, 2017; Portwood, 2005) were rated low risk for this criterion. The use of clustered RCTs made it easier for several studies to differentiate between treatment and control schools (Brigman, 2003; Curran, 2018; Haynes, 1990; Miller, 2017). Dougherty (2017) was considered low risk as the intervention groups were clearly defined, so there was no misclassification of participants. DuPuis (2013) clearly classified those receiving the intervention from those who were not through the use of a waitlist control group. Eggert (1990) recruited students into the intervention condition on a first-come, first-served basis with the remaining students being placed in the control condition, and then matched with a treatment student. Felner (1982) and Portwood (2005) clearly delineated the intervention and control students through the use of matched group design.

4.2.2.3 Selection of participants into the study (selection bias)

Low risk. All nine studies mentioned above were also considered low risk for selection of participants into the study. All participants were recruited into the study based on characteristics observed before the start of the interventions. Therefore, they were no observable issues with selection bias whereby participants were chosen based on characteristics that could produce spurious results.

4.2.2.4 Deviation from intended intervention

Low risk. Five of nine studies were considered to be low risk. Eggert (1900) and Brigman (2003) reported their implementation procedures, which indicated that the interventions were delivered as intended. Curran (2018) utilised scoring guidelines to ensure implementation fidelity for the treatment group. The teachers in the Dougherty (2017) study delivered 89% of intervention lessons and three independent evaluators were used to ensure that the programme was implemented appropriately. Miller (2017) reported extensive implementation data which revealed that 90% of teachers completed all RBI activities which a high degree of adherence.

Unclear risk. Both DuPuis (2013) and Felner (1982) described their respective programme and evaluation procedures but do not report if these procedures are carried out as intended. Therefore, the review author was unable to determine if there were deviations in programme delivery.

No information. Haynes (1990) and Portwood (2005) did not report any implementation data.

4.2.2.5 Incomplete outcome data

Low risk. Studies reporting attrition rates of 0-10% were rated as low risk (Curran, 2018; Felner, 1982; Miller, 2017; Haynes, 1990). Dougherty (2017) reported a high attrition rate of 36%. However, there was no differential attrition and the authors used the full-information likelihood method to account for incomplete data. DuPuis (2013) was deemed low risk as several analyses have a degree of freedom of 78, which would indicate that they have outcome data for most of their 83 participants.

Serious risk. Eggert (1990) was given a high-risk rating on this criterion. 26% of students dropped out of the intervention group and 39% dropped out of the control group, and the researchers did not adequately address attrition in their analyses. 37.8% of students were missing academic outcome data in the Portwood (2005) study and it appears that no analyses were done to account for the high levels of incomplete data.

Critical risk. Brigman (2003) reported high levels of attrition from both intervention conditions, (treatment: 47.2%; control: 30.6%) and does not control for the effects of missing data in their analyses. Additionally, no reasons were provided for the overall or differential attrition rates. The academic achievement data included in the forest and harvest plots of this review are likely to be severely impacted by these factors, therefore this study was more critically assessed, and considered a critical risk study.

4.2.2.6 Measurement of outcomes

Low risk. There were no apparent errors in outcome measurement or differences in assessment across condition groups. Moreover, there is a low likelihood that the knowledge of the intervention could influence outcome measurement so all nine studies were considered low risk.

4.2.2.7 Selective Reporting

Selective reporting was assessed for both effectiveness and equity outcome data. As noted earlier in sections 4.1.1 and 4.2.1.6, all studies but Tak (2014) not have any published protocols with which reporting on outcomes could be compared to identify potential selective reporting. Therefore, decisions on selective reporting were left to the review author's judgements.

Low risk. Both Brigman (2003) and Miller (2017) reported data on the effectiveness and equity impact of the intervention, and so they were considered low risk. Haynes (1990) was considered low risk as the authors indicated that they did not have additional effectiveness or equity data, so selective reporting was unlikely.

Unclear risk. Though Curran (2018) provided sufficient effectiveness data, it was difficult to determine whether additional data exist on equity outcomes, so it was considered an unclear risk of bias. Both Dougherty (2017) and DuPuis (2013) report only the beta coefficients and p-values across multiple analyses instead of basic statistics commonly found in studies, like means and standard deviations. Though the analyses and measures appear to be robust, it is unclear whether or not the authors purposely chose to not report the full data required for meta-analyses. Felner (1982) also reported partial necessary data, such as means without standard deviations, but fully presented the values of F-tests, p-values, and t-tests. In terms of selective reporting of equity data, all three studies report equity-relevant information on students of varying racial, socioeconomic, and gendered backgrounds so it is unclear whether the studies did not assess outcome data by these factors or if they chose not to report them.

Eggert (1990) is missing essential data in their ANOVA tables evaluating the impact of the intervention on GPAs. As the authors did not respond to multiple requests for additional data, it cannot be determined if this reporting decisions was made intentionally. With respect to equity data, it was determined from their data collection efforts on participant characteristics that no equity data was measured or analysed. Portwood (2005) was judged to be an unclear risk of bias for reporting effectiveness data. The authors reported means and sample sizes post-intervention to support their findings, but without standard deviations, a standardised mean difference cannot be calculated to fully assess their findings. In terms of equity data, it was surmised that the researchers did not measure any equity-relevant data so it was unlikely that they selectively reported this information. Figures 8 and 9 below represent the risk of bias summary and graph for non-randomised trials, respectively.

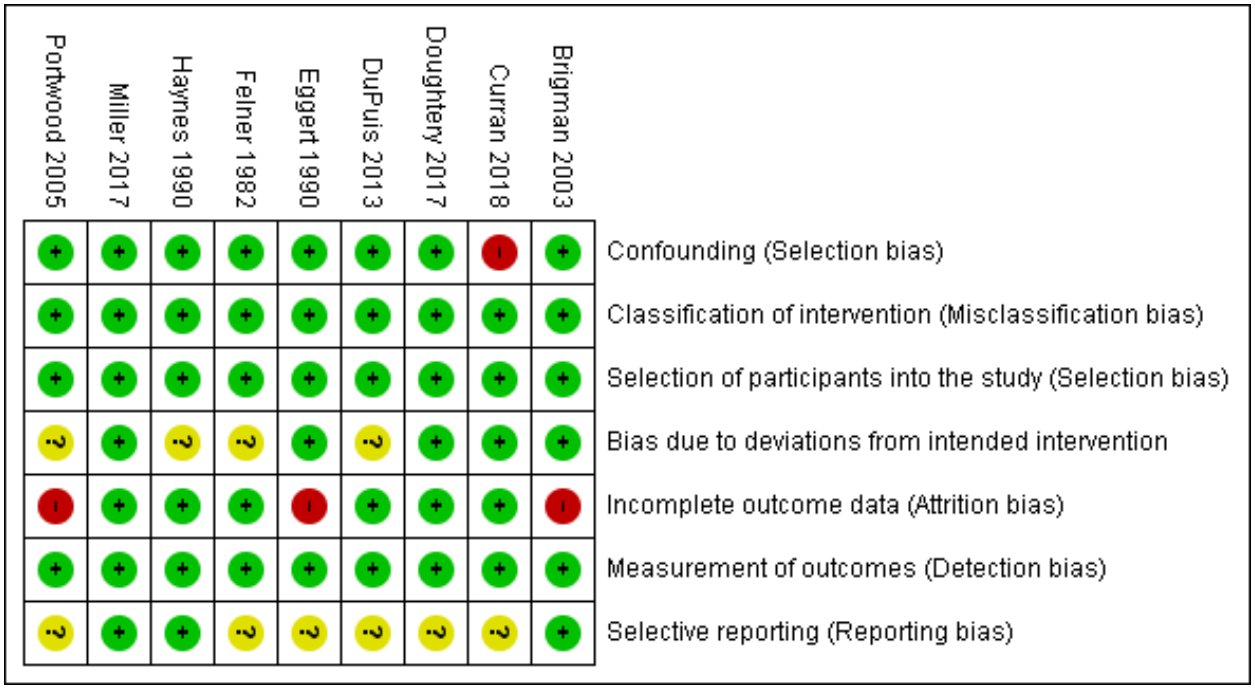


Figure 8. Risk of bias summary (ROBINS-I).

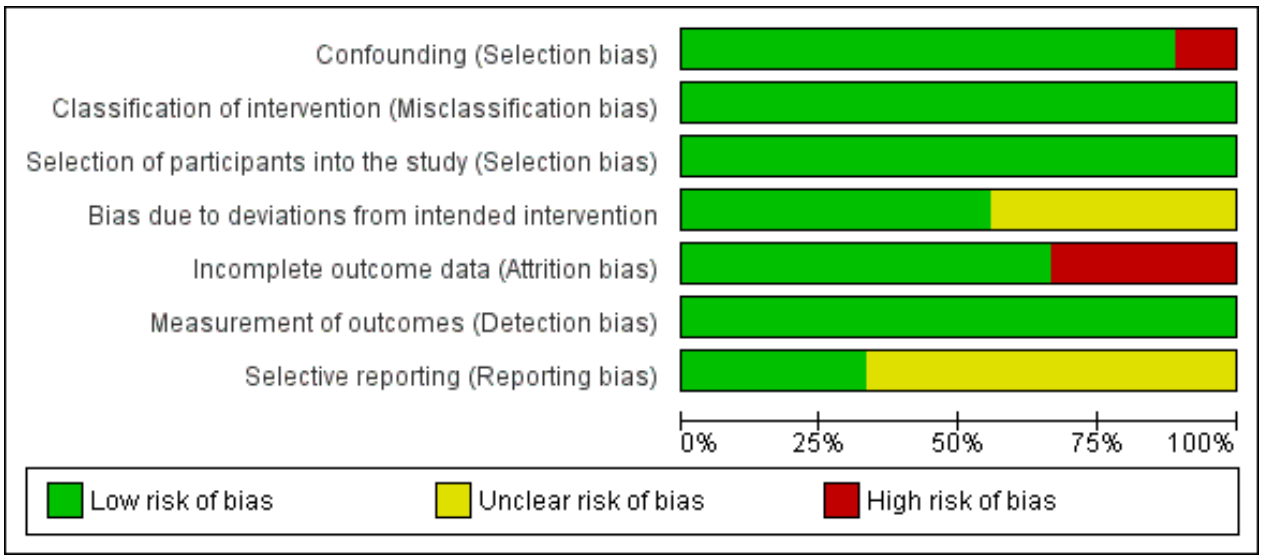


Figure 9. Risk of bias graph (ROBINS-I).

4.2.3 Overall risk of bias rating

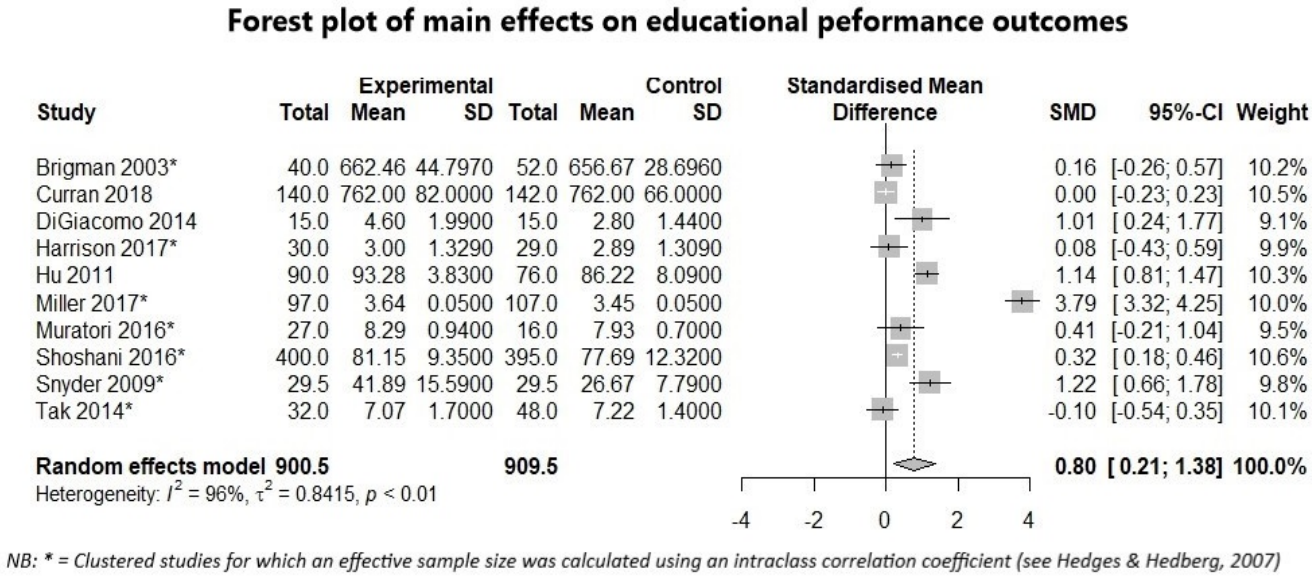
Of the 15 experimental studies, only Muratori (2016) was rated as low-risk overall. Arif (2017), Asmus (2017), DiGiacomo (2014), Eskreis-Winkler (2016) Harrison (2017), Hu (2011) Paunesku (2015), and Snyder (2009) were all deemed as some concerns. Finally, Cohen (2006), Karcher (2008), Reyes (1994), Shoshani (2016), Tak (2014), and Terry (2016) were considered as high risk. For the nine quasi-experimental studies, only Miller (2017) was considered as low risk overall. Dougherty (2017), DuPuis (2013), Felner (1982), and Haynes (1990) were deemed moderate risk (represented as unclear risk on the graphs above). Serious and critical risk studies (designated as high risk in section 3.4.4) are Brigman (2003), Curran (2018), Eggert (1990), and Portwood (2005).

4.3 Effects of interventions

4.3.1 Main findings

Out of the 24 included studies, only ten studies were used in the meta-analysis of main effects due to issues with reporting practices. These studies were used to quantitatively address the research question: **Are school-based psychosocial interventions which target academic-related protective factors compared to treatment as usual effective in improving educational performance among primary and secondary students?**

The result of the meta-analysis indicated a large, positive effect of these interventions on educational performance (g=0.8, SE=0.43; p=0.008, 95% CI = 0.21, 1.38). Additionally, the random effects model presented considerable statistical heterogeneity across all metrics (I^2=96%; T=0.84; Q(df=9)=253.59, p<0.01) (Deeks, 2011) (see figure 10).



*NB: * = Clustered studies for which an effective sample size was calculated using an intraclass correlation coefficient (see Hedges & Hedberg, 2007)*

Figure 10. Forest plot of intervention effectiveness on academic performance outcomes.

Across the different outcomes of GPA, standardised and individualised test scores, and course-specific marks, eight of the 10 included studies had a positive effect on educational performance (Brigman, 2003; DiGiacomo, 2014; Harrison, 2017; Hu, 2011; Miller, 2017; Muratori, 2016; Shoshani, 2016; and Snyder, 2009). Of the eight studies, DiGiacomo (2014), Hull (2011), Miller (2017), and Snyder (2009) have large effect sizes, i.e. $g \geq 0.8$ (Cohen, 1988). Muratori (2016) and Shoshani (2016) have medium effect sizes (Hedges' *g* approximately around 0.5) and Harrison (2017) has a small effect size, i.e. $g \leq 0.2$. One study reported no effect of the intervention on educational performance (Curran, 2018), and Tak (2014)'s study had a small, negative impact on academic outcomes. Five studies had a statistically significant effect (DiGiacomo, 2014; Hu, 2011; Miller, 2017; Shoshani, 2016; and Snyder, 2009) whereas the other five did not as their confidence values include zero, which is the value of no effect (Brigman, 2003; Curran, 2018; Harrison, 2017; Muratori, 2016; and Tak, 2014).

4.3.2 Robustness checks

4.3.2.1 Sensitivity analysis by risk of bias rating

Given the varying levels of study quality, a sensitivity analysis was conducted according to risk of bias ratings. Out of the 10 studies used in the meta-analysis of main effects, only Muratori (2016) and Miller (2017) were included in the analysis as they were considered to be low risk overall. This resulted in Brigman (2003), Curran (2018), DiGiacomo (2014), Harrison (2017), Hu (2011), Shoshani (2016), Snyder (2009), and Tak (2014) being excluded. Thus, caution should be used when interpreting the effectiveness findings given the study quality for the review's body of evidence. When accounting for only high-quality studies, the sensitivity analysis produced a much larger effect (g =2.11, SE=1.69, p=0.19, 95% CI =-1.20, 5.41) in comparison to the meta-analysis of intervention

effectiveness (g=0.8, SE=0.43; p=0.008, 95% CI=0.21, 1.38). However, the effect estimate of the sensitivity analysis was statistically insignificant. Additionally, the random effects model indicated that there was considerable heterogeneity (I^2=99%; T=5.61; Q(df=1)=72.47, p< .0001), which increased from the I^2 and Tau metrics in the model of main effects (I^2=96%; T=0.84; Q(df=9)=253.59, p< 0.01) (Deeks, 2011). Figure 11 represents the forest plot for the sensitivity analysis by risk of bias.

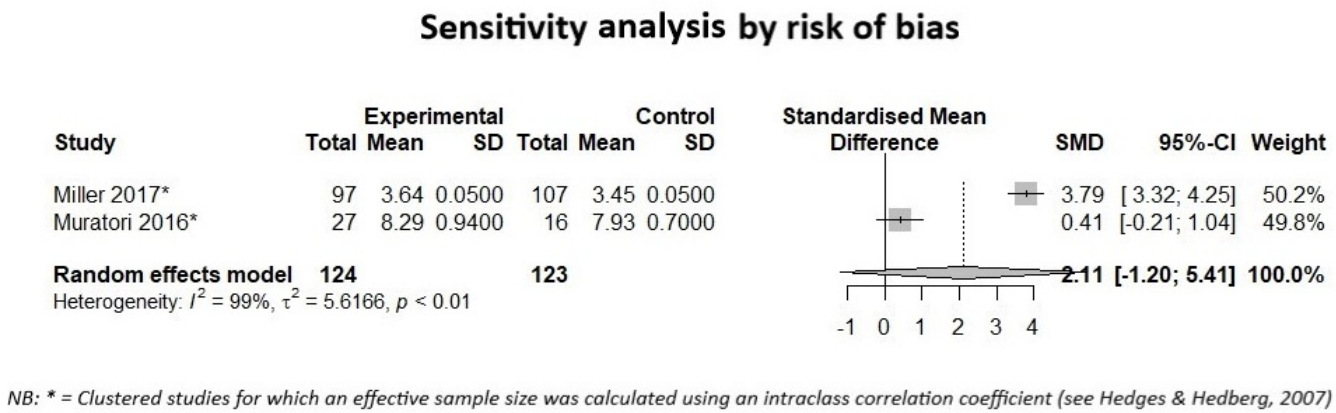


Figure 11. Forest plot of sensitivity analysis for low-risk studies. Miller (2017) was rated using ROBINS-I and Muratori (2016) was rated using RoB 2.0.

4.3.2.2 Checks for heterogeneity using influence diagnostics

Given the high levels of heterogeneity in both meta analyses, influence diagnostics were performed as an additional robustness check. This allowed the review author to identify which study or studies may be contributing the most to the between-study variability. The results from the leave-one-out meta-analysis suggest that both Miller (2017) and Muratori (2016) had the greatest influence on the degree of heterogeneity observed. Omitting Miller (2016) led to no heterogeneity across the remaining studies (I^2=0.00; T^2=0.00; Q(df =9)=0.25, p=0.54). Moreover, when Muratori (2016) is removed from the model, the statistical heterogeneity substantially decreases from the original model of main

effects (I^2=39.45%; T^2=3.22; Q(df=9)=9.46, p=0.11 vs. I^2=96%; T=0.84; Q(df=9)=253.59, p<0.01). The respective amounts of heterogeneity attributable to the remaining studies are similar to the heterogeneity metrics observed in the meta-analysis of main effect, and as such, will not be reported.

4.3.2.3 Checks for bias using funnel plots and Egger's test

A final robustness check was conducted using a funnel plot and an Egger's test to account for possible bias in the results of intervention effectiveness. The funnel plot visually presents asymmetry by study effect size. Additionally, the Egger's test statistically summarises any bias, which could be the result of possible reporting bias, publication bias, and/or statistical heterogeneity (Egger et al., 1997). Upon visual inspection of the funnel plot, there appears to be asymmetry in the findings (see figure 12 below). Six of the ten studies are unevenly populated to the left of the grand mean (g=0.8), and Miller (2017) serves as an outlier given its very large effect size. If there was little to no bias present in the body of evidence, then the number of studies as well as the positioning around the 95% confidence interval lines should be balanced, which is not the case with these studies. These findings are supported by the results of Egger's test. There was a large amount of bias (estimate: 3.52; SE=3.26) which suggests that the null hypothesis of funnel plot symmetry could not be supported (t=1.0789; df=8; p=0.31). Taken together with the findings of the Egger's test, it can be concluded the asymmetry observed in the funnel plot is due to bias, which is attributable in large part to between-study variability.

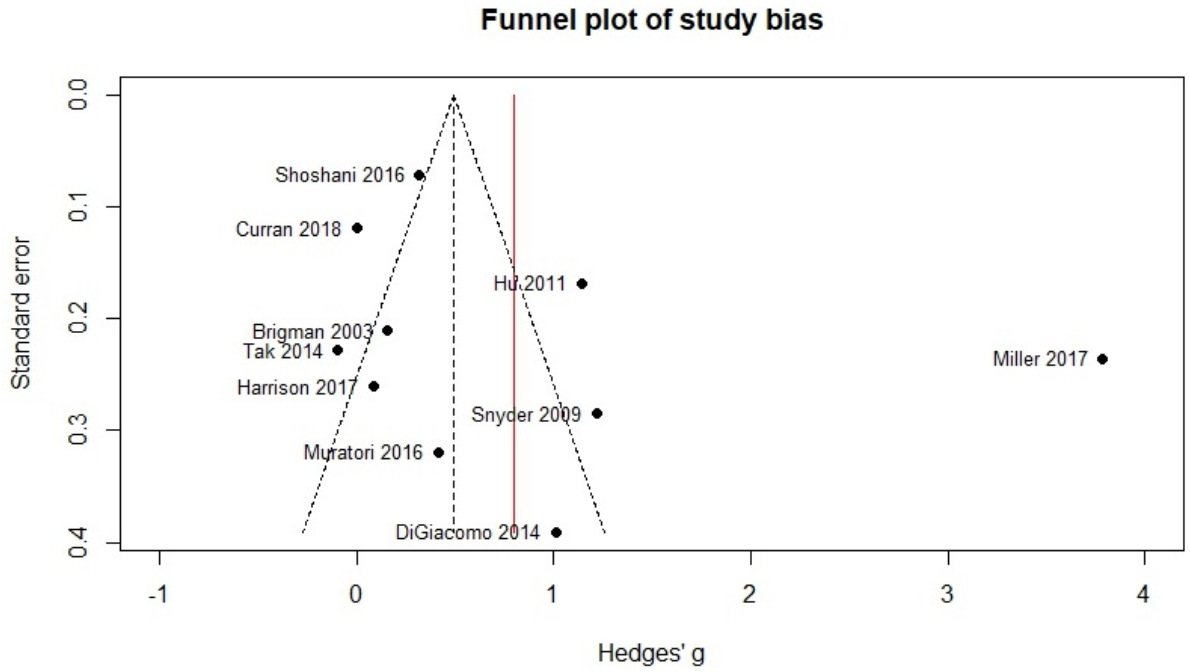


Figure 12. Funnel plot of asymmetry to assess study bias(es).

4.4 Equity effects of interventions

4.4.1 Description of equity evidence and equity-relevant information

Equity evidence. As noted in the methods portion of the review (see section 3.4.2.2.), equity evidence refers to interaction and moderator analyses that statistically test for differential effects of the intervention by PROGRESS-Plus factor. Only four studies reported equity evidence (Brigman, 2003; Cohen, 2006; Eskreis-Winkler, 2016; Miller, 2017). The Brigman (2003) study reported equity data on race in their 2007 report by conducting a MANCOVA comparing all races on standardised math test scores (Miranda et al., 2007). An additional MANCOVA was reported for the effect of the intervention by race and treatment condition. Cohen (2006) reported interactions of treatment condition by both gender and race using multiple regressions. Both Eskreis-Winkler (2016) and Miller (2017) reported subgroup effects by race and interaction effects of gender by

condition using hierarchical linear regression models. No studies provided any equity data on the basis of socioeconomic status, so this section was left blank on the harvest plot.

Equity-relevant data. Unlike equity evidence, equity-relevant data do not analyse the intervention effect by equity factors. Instead they provide information on the presence of PROGRESS-Plus factors in a given study population. In regards to race and ethnicity, most studies (n=18) reported baseline sociodemographic information. Only two studies reported adjusted associations. Eskreis-Winkler (2016) adjusted post-intervention performance outcomes by race given that there were more white students in the control group relative to the treatment group. Cohen (2006) reported covariate-adjusted means which accounted for race and performance level for White and African-American students. Adjusting the data based on equity factors in this case accounts for the significance of race in the outcomes, which demonstrates the extent to which the intervention may have an impact on students from differing backgrounds. For gender, all but four studies presented sociodemographic information (n=20), and none reported adjusted associations. Fourteen studies reported on socioeconomic status, with no data on adjusted associations. In total, there were 52 counts of reported baseline characteristics across the PROGRESS-Plus factors. Cohen (2006) and Eskreis-Winkler (2016) were the only studies that reported adjusted associations of any kind. Figure 13 below displays both equity evidence and equity-related information availability.

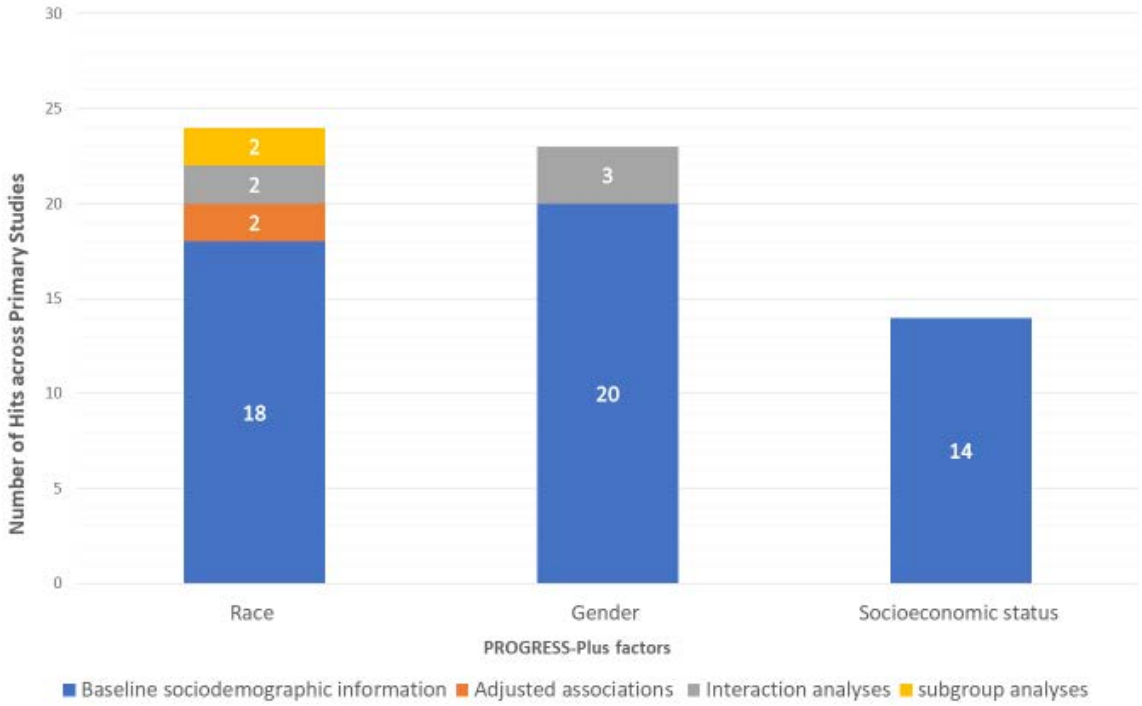


Figure 13. Distribution of equity evidence and equity-relevant data across the included studies.

4.4.2 Acknowledgement of equity

Both Cohen (2006) and Brigman (2003) (as reported in Miranda et al, 2007) acknowledged equity objectives as these studies sought to reduce the achievement gap between White and Black students in their interventions. Although three of the four studies included in the harvest plot stated that they performed their moderator analyses *a priori* (Brigman, 2003; Cohen, 2006; Miller, 2017), none of them specified that these analyses were done in the effort of reducing educational inequity.

4.4.3 Main findings of equity data synthesis

Of the 24 studies included in this review, only four studies provided enough evidence to quantitatively address the following research question: **Do school-based psychosocial interventions fostering academic resilience have equitable effects on academic**

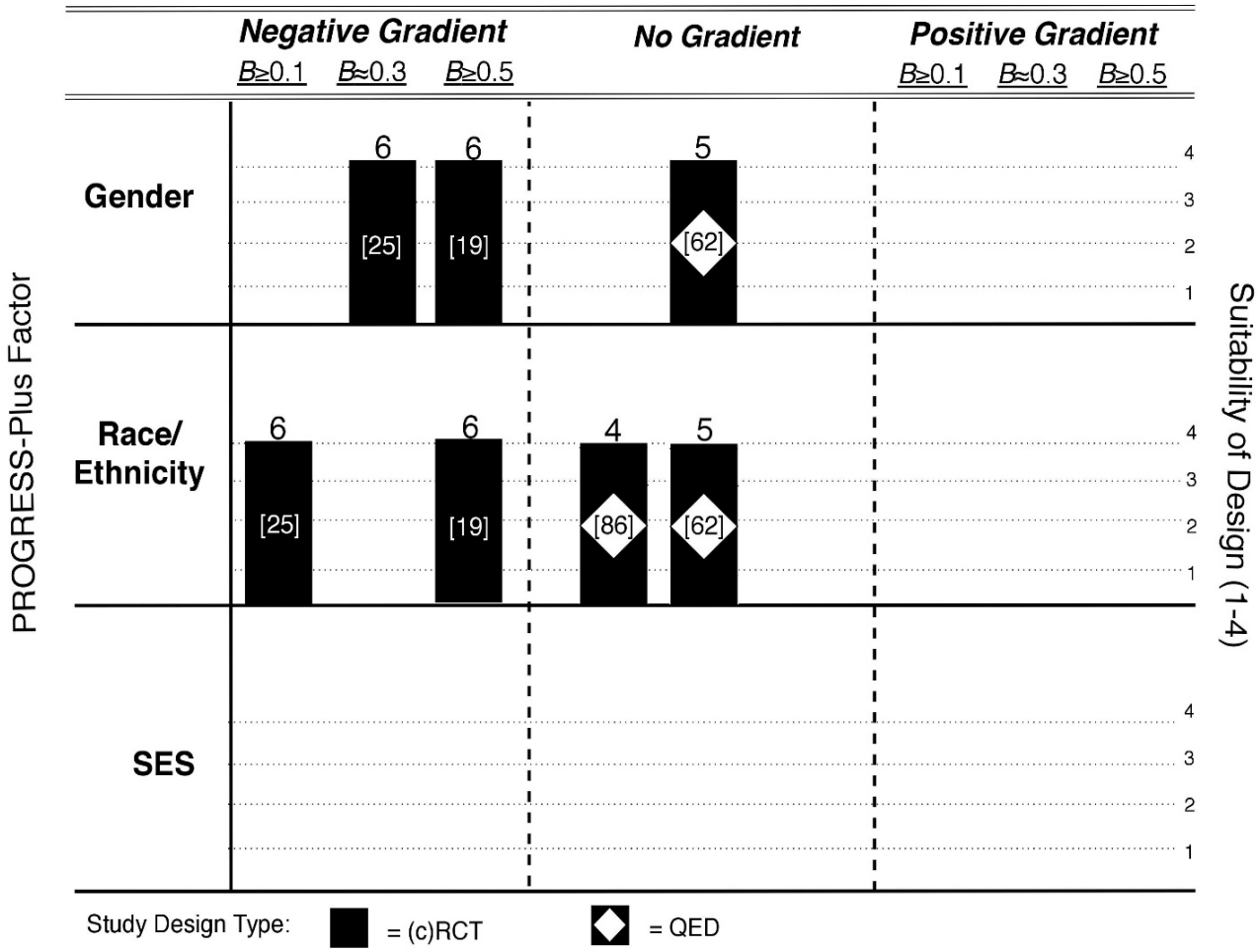
gains for factors of race, gender, and socioeconomic status among primary and secondary students?

Gender. The results from the harvest plot below indicate that the interventions favour disadvantaged students (i.e. the negative hypothesis) for the PROGRESS-Plus factor of gender. Referring back to the introduction, gender inequity in education is very context-specific and varies greatly depending on a number of factors (see section 1.2.4). In regards to maths, research suggests that they may be a gendered effect such that girls may underperform relative to the male peers due to stereotype threat (Beilock et al., 2010; OECD, 2012). Therefore, since Cohen (2006) and Eskreis-Winkler (2016) report math-related academic outcome data, female students were considered as the 'disadvantaged' group. The results from the linear regression in Eskreis-Winkler (2016) indicate an interaction effect such that female treatment students receive benefit from the intervention in comparison to male students (B=0.18; SE=.14; p>0.05).

Cohen (2006)'s interaction effects from their linear regression suggest a large, positive effect of the intervention on black female students in the treatment condition (B =.64; $t(97)$=2.21; p=0.03). The intervention had a negative effect on white female treatment students (B=-.48; $t(51)$=-2.49; p< 0.02) and no effect on all male treatment students ($|B's|$<0.16; $|t's|$-1.1; p=N.S.). The authors suggest that the intervention may be more beneficial to disadvantaged students like black female students in a context where negative stereotypes on both gender and race could occur, such as in an advanced math class. Miller (2017) found no differential intervention effects by gender (B =-.03; SE=.06, p=N.S.). Comparing effect estimates demonstrate that the magnitude of effect from both Cohen (2006) and Eskreis-Winkler (2016) are larger and statistically significant relative to the effect reported in Miller (2017), which suggest that the direction of effect overall favours the 'disadvantaged' group, female students.

Race. With respect to equity effects by race, two studies (Cohen, 2006; Eskreis-Winkler, 2016) favour the disadvantaged students (i.e. non-Caucasian students) and two studies benefit neither disadvantaged (i.e. non-Caucasian) nor advantaged (i.e. Caucasian) students (Brigman, 2003; Miller, 2017). The regression model reported in Cohen (2006)'s online supplementary material noted a significant interaction effect for African-American students (B=.20; t(109)=2.32; p< 0.02). This finding was supported by the fact that African-American students in the treatment group were the only group whose GPA did not deteriorate during the duration of the intervention in comparison to black students in the control group, and white students across both treatment and control conditions ($_M$decline = -0.25, -0.19, and -0.24, respectively). For Eskreis-Winkler (2016), the moderator analyses suggest a positive impact on both Asian and Hispanic students (B=0.76, SE=0.3, p<0.05; B=0.16, SE=0.22, p>0.05, respectively), though the programme produced a small, negative effect among African-American students (B=-0.09; SE=0.2; p>0.05).

Brigman (2003)'s one-way MANCOVA found essentially no difference in intervention impact among students of varying ethnicities in terms of interaction analyses (η^2=0.04; Wilks' λ =.992; F [4, 2228]=2.36; p >.05). This was also true of the race by treatment condition interaction effect (η^2=0.04; Wilks' λ=.999; F [4, 2228]=.258; p >.05). Both results suggest treatment students' educational performance improved irrespective of race (see section 4.3.1 for effectiveness results). Miller (2017)'s multilevel regression model yielded a negligible effect for student ethnicity as well (B=0.01; SE=0.03, p=N.S.). Comparing the results by race, the findings from Cohen (2006) and Eskreis-Winkler (2016) may outweigh the null effects from Brigman (2003) and Miller (2017), suggest that the direction of effect overall favours the 'disadvantaged' group, minority students. Figure 14 below displays the



harvest plot of equity effects by PROGRESS-Plus factors, with the SES section left blank as no equity evidence was reported on this characteristic.

Figure 14. Harvest plot of educational performance by PROGRESS-Plus factor. On the left side, there are the 3 PROGRESS-Plus factors of interest (gender, race, and socioeconomic status). Ear bar represents the equitable impact of the intervention, sorted by its effect on a study's disadvantaged student population, (negative gradient) advantaged student population (positive gradient), or neither (no gradient). Bar colour, black, represents objective outcome data metrics, and the number on top is the methodological quality (1 being the lowest and 6 being the highest). The bracketed number is the study ID, and the non-randomised trials are demarcated by the inserted diamond figure. Effect size heuristics have been provided to demonstrate the magnitude of the intervention effect on equity outcomes. Study IDs: 19 = Cohen (2006), 25 = Eskreis-Winkler (2016), 62 = Miller (2017), and 86 = Brigman (2003).

4.4.4 Subgroup analyses

As noted in methods section 3.4.10, the review author intended to conduct between-trial subgroup analyses to statistically synthesise the findings across equity factors to complement the harvest plot results. Akin to the meta-analysis and harvest plot, lack of complete outcome data limited the ability to perform moderator analyses. Given the substandard reporting practices observed in this body of evidence, only a few studies reported the means, standard deviations, and sample sizes required for a meta-analysis. Thus, conducting a between-trial moderator analysis by race was infeasible as studies that could have been classified as 'non-White' did not report usable data. As a result, essentially all but two studies that did report academic outcome data could have been classified as White given the proportion of Caucasian students relative to other races. This would have greatly skewed the between-trial moderator analysis given the underrepresentation of different ethnicities from other included studies. In regards to gender, most studies that did provide academic outcome data were evenly split between the two genders, thus studies could not be classified as mostly male or female studies for meta-analysis purposes. In terms of socioeconomic status, reporting on this PROGRESS-Plus factor was inconsistent, not comprehensive, and varied greatly between studies, thus making classification by class standing infeasible. Therefore, it was not possible to do further equity data analyses without better reporting of trial-level data.

4.4.5 Quality assessment of equity evidence

In addition to assessing the visual and statistical assessment of the equity data, the information from the aforementioned studies were also evaluated across two quality indicators. On the first metric, suitability category, all studies received the highest score, 'level 4', meaning that the study designs included concurrent control groups and

prospective measurement of exposure and outcome. For the second indicator, methodological quality, both Cohen (2006) and Eskreis-Winkler (2016) received the highest score of 6. The studies earned a point for each of the value domains: random sampling, randomisation procedures, controlling for confounders (comparability), valid data collection instruments, low attrition rate, and causal inference (attributability to intervention). Miller (2017) received a score of 5 as this study did not use randomisation given that it was a quasi-experimental design. Brigman (2003) received a 4 as this study also did not use randomisation as it was a quasi-experimental design, and also had a high attrition rate (refer back to section 3.4.3.2 for in-depth explanations of each criterion).

In addition to the quality assessment parameters reported in the harvest plot, the review author also considered the quality of the analyses that produced the interaction and moderator effects. No protocols or pre-trial registration documents in which they specified their analytical methods could be located. However, only Eskreis-Winkler (2016) did not make an *a priori* specification regarding their moderator analyses, but it is possible that all authors could have performed these analyses post-hoc as there is no way of verifying given the lack of published protocols. This then could suggest selective reporting, which would bias the results. The two studies that clearly noted their objectives of assessing equity impacts, Brigman (2003) and Cohen (2006), do not detail how they went about analysing their data beyond noting the use of MANCOVA and multiple linear regression models, respectively. Similarly, Eskreis-Winkler (2016) did not discuss equity assessment methods. Miller (2017) has a thorough explanation of their data analysis methodology for effectiveness data but they do not explain how these decisions relate to their investigation of equity. Brigman (2003), Cohen (2006), and Miller (2017) all appear to be sufficiently powered for their moderator analyses due to their large sample sizes, although none of the studies performed power analyses to determine the minimum sample sizes required for

their analyses (Cohen, 1992). Eskreis-Winkler (2016) appears to be underpowered due to its sample size, which lowers the quality of the findings.

5 Discussion

This review aimed to determine if school-based interventions which promote academic protective factors could improve educational performance through the causal mechanism of academic resilience. This review also endeavoured to assess whether improvements in educational performance were equitable among students whose backgrounds may put them at an academic risk due to their gender, race, and/or socioeconomic status. In regards to the question of effectiveness, the results indicate that these interventions were largely effective in improving academic outcomes. However, given the use of data from high risk studies in the main effects meta-analysis, as well as the lack of statistical significance in the sensitivity analysis, these findings should be interpreted with caution. Regarding equity, the findings suggested that the interventions may decrease gender inequity among female students and racial inequity among minority students, but there were no findings available for equity impacts by socioeconomic status. The following section discusses the interpretations of both the effectiveness and equity findings and their related conceptual and methodological considerations. This section also places a particular emphasis on the external validity of the results by assessing the overall quality and applicability of evidence.

5.1 Effectiveness

5.1.1 Interpreting effectiveness findings

The findings suggest that academic resilience may be a worthwhile and innovative psychosocial tool for improving scholastic abilities, which supports the review author's

assertions that academic resilience is not only the result of early life experiences and traits, but rather, a skill that can be developed throughout a student's life. The treatment effect was large (g=0.8) and statistically significant (p=0.008) and this value could be attributed to a number of causes. Looking back on the review's theory of change, by fostering psychosocial protective factors, students may have learned academic-related metacognitive strategies, which reinforced positive academic behaviours, leading to better school performance. This is supported by research that suggests that these skills can increase academic achievement in formal schooling (McClelland, Acock, & Morrison, 2006). Other causal pathways could include the promotion of positive behaviours which offset disruptive behaviours known to impede learning in school environments (Flay & Allred, 2010). The results could also be associated with implementation effort as programme fidelity has been found to be a vital aspect of intervention success (Durlak & DuPre, 2008). Interventions that were delivered regularly over a longer period of time may have led to greater intervention impact, which could suggest a dose-response effect. For example, Snyder (2009) noted that their programme may have had such a large effect because it was delivered weekly for multiple school years. Moreover, several interventions used interactive methods that may have reinforced the intervention delivery by creating enriching interactions with their teachers and peers, thus increasing its effect.

The effectiveness findings could be also the result of factors not directly related to the intervention and for these reasons, they should be interpreted with caution. Namely, performance bias may have elicited a Hawthorne effect, which refers to the reactive behaviour of students who may be aware that they are being observed. (Adair, 1984). There is a chance, though likely very slim, that the presence of differential treatment irrespective of the school-based programme could have influenced students' performance. Lastly, though the interventions were limited to environmental and dispositional protective factors

that target the individual and the school setting, there could be a myriad of factors that improved school performance, including the home environment. This factor has the single largest impact on a child's academic development outside of the school setting (Walberg, 1986). Though the home environment was beyond of the scope of this review given the focus on school-based programmes, it is still important to acknowledge its potential effect on the outcomes.

5.1.2 Findings of intervention ineffectiveness

Curran (2018) and Tak (2014) were the only two studies in the meta-analysis that had a null effect and a small negative effect, respectively. Given these results, it is important to consider how they may have differed from the other studies. Tak (2014) may have caused marginal educational decline because the universal intervention exposed students who were not at risk of academic failure to iatrogenic effects. Students may have been alerted to their own negative cognitions and shortcomings in coping abilities through the CBT intervention, making them feel less capable in school which led to academic downturn. The authors also note potential issues with recall bias. Students were asked to self-report their most recent grades by recall, so the reliance on self-report data could have caused inaccuracies in academic outcome measurement.

Regarding Curran (2018), the null effect is likely the result of poor implementation due to insufficient personnel training. Several teachers noted that they were not properly trained in how to deliver the intervention. Since these teachers also taught both control and intervention classes, their methods of teaching may have been so similar given their inability to effectively implement the intervention, it would have been difficult to detect a treatment effect. However, evidence of no effect and no evidence of effect should not be conflated (Schünemann et al., 2008). The intervention utilised in the Curran (2018) study

may produce positive effects in different contexts and thus more research may be needed before making a clear judgement.

5.1.3 Conceptual and methodological considerations of effectiveness data

Methodological. The effectiveness findings reflect a wide array of interventions, data collection and delivery methods, various types of outcome indicators (i.e. GPA, test scores, and course-specific marks), and differing outcome metrics, be it objective or self-reported data. Given the varied methodological tools and the limited number of studies that provided proper data, it was not possible to investigate whether one type of intervention, delivery method, or outcome type was the most appropriate. Additional research could elucidate which methods are the most effective concerning academic resilience interventions.

Conceptual. One must also consider how 'educational performance' was conceptualised to shed light on how the findings of effectiveness should be interpreted. There is no standardised way in which academic achievement, and thus related measures of performance, are presented in the body of evidence. Some authors refer to academic performance in reference to a one-semester class, such as Chinese or maths, whereas other authors represent academic achievement as performance averaged across several classes over time. As such, these findings on effectiveness as they pertain to academic gains may reflect differing parameters of what is meant by improving one's scholastic abilities. As alluded to above, a number of factors could have led to improved academic performance. Though subjective well-being outcomes were outside of the scope of this review, testing the causal mechanisms of psychosocial functioning as a result of academic resilience to better understand comparable gains in educational performance could possibly reinforce the validity of these findings.

5.2 Equity

5.2.1 Interpreting equity findings

The equity findings suggest that interventions may improve equity gaps in education for female students and minority students as the direction and magnitude of effect in the harvest plot favours the 'disadvantaged' groups. This is an important finding as it suggests that these interventions tend to decrease inequity. This helps to rectify some of the sources of gender and racial inequity outlined in the introduction, such as stereotype threat. As is the nature with most reports, very few included studies reported any equity evidence which investigated programme effects among disadvantaged populations (Welch et al., 2015). Only one study, Cohen (2006), explicitly sought out to reduce inequity gaps between the most and least advantaged students (i.e. white and black students) and to reduce inequity gradients (i.e. improving academic performance among African American students of various risk levels). There was a clear absence of socioeconomic equity evidence as no studies reported any subgroup effects or interaction effects. The lack of equity evidence on this PROGRESS-Plus factor may be attributable in part to the fact that very few studies report interaction and moderator effects. If these data are reported, they typically relate to characteristics such as race and gender, so there is a dearth of evidence in the current research on equitable effects by socioeconomic status.

While virtually all studies reported equity-relevant information across the PROGRESS-Plus factors, almost none reported detailed baseline demographic information, especially on socioeconomic status. Since socioeconomic status was not reported comprehensively, it stands to reason that it would be seldomly analysed as well, thus limiting the likelihood of applicable harvest plot data. Equity-relevant data are not as insightful as equity evidence in determining improvements in educational inequity. Rather,

equity-relevant data signal if the results would be applicable to certain populations if they are represented in the study sample of a given trial. Additionally, as alluded to above, sufficient equity-relevant data point towards the likelihood of further analyses that assess effectiveness among different sociodemographic groups of interest (Welch et al., 2010). Therefore, the overall lack of socioeconomic data suggests research implications for future equity research to address how to improve the impacts of poverty on educational performance.

5.2.2 Methodological and conceptual considerations of equity data

In addition to limited equity evidence, most studies do not even discuss why they do not provide this type of information. To better understand the scarcity of equity evidence, one must consider both the conceptual and methodological limitations of conducting and evaluating equity data.

Methodological. Though moderator analyses between trials were not possible due to poor reporting practices, it is nevertheless important to discuss their utility in equity reviews as they pertain to harvest plot data. The credibility of subgroup effects, or lack thereof, has incited an ongoing, intense academic debate (Oxman & Guyatt, 2002; Petticrew et al., 2011; Sun, Briel, Walter, & Guyatt, 2010). Some researchers dispute the utility of subgroup analyses as they may produce spurious results (Oxman & Guyatt, 2002). This stems from the methodological issues of conducting subgroup analyses, such as underpowered samples which lead to inadequate statistical power, thus weakening the validity of the moderator effects. There are also specific criteria which restrict a researcher's ability to perform a subgroup analysis. This includes having a clear, theoretical justification for conducting a moderator analysis, *a priori* specification and direction of impact, and baseline measurement of the variables of interest (Oxman & Guyatt, 1992; Sun et al., 2010).

If the necessary prerequisites are not met and explicitly noted in the research, then subgroup effects may be considered invalid or, perhaps worse, susceptible to p-hacking.

With respect to the current research, three of the four studies included in the harvest plot specified their subgroup and interaction analyses *a priori* (Cohen, 2006; Brigman, 2003; Miller, 2017). Alternatively, some researchers support them as worthwhile evidence for policy and practice frameworks (Petticrew, 2011). Both moderator and interaction effects suggest differential benefit, or harm, of school-based interventions which have significant implications on how these programmes can be applied to certain student populations. Especially in the field of educational equity research, these data are vital to understanding if academic resilience interventions are indeed effective at reducing inequity in school performance.

Though the harvest plot suggests that academic resilience interventions may be effective in reducing educational inequity, this is based on subgroup and interaction effects which may lack statistical rigor. Further, harvest plots are highly useful tools for examining the distributional effects of equity. However, they are only visual in nature and as such, do not use statistical synthesis to quantify impact, which leaves judgements on equity effects open to the researcher's discretion. To rectify this issue given the lack of moderator analyses, the review author included effect size heuristics to demonstrate the magnitude of effect. Contextualising intervention effectiveness across PROGRESS-Plus factors helps in determining the extent of their utility. This, however, is a minor solution to a larger problem of improving certainty in the interpretation of equity findings. Therefore, in light of these considerations, the equity evidence presented in this review should be viewed positively, but with cautious optimism.

Conceptual. Additional concerns that warrant consideration is how 'disadvantage' is conceptualised. This was a common issue in regards to socioeconomic status, as most study authors did not provide specific definitions of how social standing was operationalised in their research if they reported socioeconomic data at all. This conceptual limitation is important given the global purview of this review. Students who would be considered as 'low-income' in a high-income country like the Netherlands in Tak (2014)'s study may equate to 'higher income' in Arif (2017)'s research in Pakistan, which is a lower-middle income country (World Bank, 2016). Without clear conceptualisations of socioeconomic status, it becomes difficult to standardise and thus compare inequities across studies from different countries. Though race may be relatively more consistent in terms of inequity for those considered as minority groups, the effects of educational disadvantage on gender is very context-specific. As discussed in the introduction, gender interplays with geographical location, race, and global trends. Using the between-country comparison once more, the female students in Harrison (2017) could be considered as higher risk relative to the female students in the Muratori (2016; Italy) study given the cultural preference towards educating boys, especially in rural China where this study is based (Wang, 2005). Additionally, all of the students in the Harrison (2017) study are from low-income backgrounds, which further decreases the likelihood of girls receiving a worthwhile education due to structural barriers for poor female students (UNESCO, 2013a). These conceptual considerations are not fully addressed in the body of evidence that informs this review, which may be another reason as to why there is such limited equity evidence available.

5.3 Overall completeness of evidence

5.3.1 Overall completeness of evidence

Effectiveness. The twenty-four studies identified for this review are believed to be mostly representative of the scope of school-based interventions for academic resilience which have been trialled and this strengthens the external validity of the findings. This is especially pertinent among the type of participants as the review had no geographic restrictions. Thus, the review's study population encompassed students from several countries and cultural backgrounds, which suggests that these interventions may be effective across a wide variety of academic settings. Moreover, the review author attempted to investigate all pertinent interventions to improve the generalisability of the findings by identifying as many relevant studies as possible. This resulted in seven protective factors being represented in the body of evidence, which are academic motivation, academic planning, academic perseverance, academic self-efficacy and self-regulation, coping, cognitive style, and school-based social support, though evidence for two protective factors could not be located (i.e. adaptive help-seeking and self-care behaviours).

Despite the lack of published protocols and inadequate reporting standards which may suggest selective reporting, two-thirds of the protective factors were included in the meta-analysis (i.e. academic motivation, academic perseverance, cognitive style, coping, school-based social support, and academic self-efficacy and self-regulation), which bolstered the generalisability of the statistical evidence. Finally, all relevant academic outcome measures were considered. This is an important consideration because the findings suggest that these interventions may ameliorate academic performance in terms of GPA, individual marks, standardised test scores, and numerous other ways.

Equity. In terms of the comprehensiveness of the equity evidence and the generalisability of its findings, multiple attempts were made by the review author to secure all possible equity evidence for the data synthesis. The authors of ten studies could not provide additional data for the harvest plot (see appendix 7.4), and Brigman and colleagues did provide additional equity information. Three studies had sufficient data already reported (Cohen, 2006; Eskreis-Winkler, 2016; Miller, 2017) and seven studies were determined as having no additional equity information (DiGiacomo, 2014; Haynes, 1990; Harrison, 2017; Hu, 2011; Muratori, 2016; Shoshani, 2016; Terry, 2016). Therefore, 21 studies were accounted for in securing evidence on equity impacts, so within the scope of this review, the equity data are considered to as complete as possible.

5.3.2 Applicability of findings and explaining heterogeneity of evidence

Though this review has reported the most completed evidence possible, one must consider if the evidence is also applicable, as applicability of the findings is another vital part of determining external validity (Green & Glasgow, 2006). The applicability of the evidence can be determined by assessing the impact of heterogeneity. Given the high level of statistical heterogeneity in the meta-analysis of main effects, several robustness checks were performed as explanatory analyses to highlight any factor(s) that may compromise the validity of the results. The large amounts of variation observed in the dataset is likely due to methodological heterogeneity because of the different types of programme components, delivery methods, and measurement tools used across the studies. This is a plausible explanation for Miller (2017) which is accountable for 100% of the heterogeneity according to the influence diagnostics. Of the 10 studies included in the meta-analysis, this study was the most wide-ranging in terms of intervention content. It delivered 21 activities over five units which were seemingly disparate (i.e. diversity and inclusion, empathy and critical

thinking, communication, problem solving, and peer relationships). Some of these units, such as diversity and inclusion and empathy and critical thinking, were more closely related to social and emotional learning interventions than to those of academic resilience. However, this study was included as the intervention was considered to be more than 50% applicable to academic resilience. However, the inclusion of SEL components may have added to the between-study variability since these components could not be separated out. This conclusion is also corroborated by the results from the funnel plot, in which Miller (2017) is the only outlier. Because this study is a massive outlier, its impact on the overall study bias is large due to the amount of heterogeneity this study introduces. Having identified Miller (2017) as the main contributor to the overall bias and heterogeneity, the review author believes that the findings of both the effectiveness and equity portions of this review are applicable and externally valid.

5.4 Quality of evidence

The corpus of studies included in this review presented some concerns about the quality of the evidence. This is due in large part to the uncertainty if vital evidence was purposely excluded because it did not support the study authors' desired findings. Only Miller (2017) and Muratori (2016) were considered low risk of bias and were included the sensitivity analysis of study quality. The results show that the treatment effect of the high-quality studies was substantially large (g=2.11). However, this value was statistically insignificant (p=0.19) in addition to being very high in heterogeneity (I^2=99%; T=5.61; Q(df=1)=72.47, p< .0001), implying that the results based on study quality may not lend any validity to the findings. Furthermore, a total of ten studies were considered as high-risk studies, four of which were included in the meta-analysis (Brigman, 2003; Curran, 2018; Shoshani, 2016; and Tak, 2014) and two of which were synthesised in the harvest plot

(Brigman, 2003; Cohen, 2006). Given that almost half of the data used to calculate both the treatment effect and the equity effects were obtained from high-risk studies, there is the possibility that these findings could have been skewed by study bias, so these findings should be interpreted cautiously.

Finally, the inclusion of weaker non-randomised designs, like pretest-posttest studies, may have diluted the quality of the evidence. Four of 24 studies used this study design (Curran, 2018; DuPuis, 2013; Haynes 1990; and Miller 2017) which is arguably the weakest design as it typically controls for confounders with the least rigour, thus posing the greatest threat to causality. Given the issues with reporting, only Curran (2018) and Miller (2017) were included in the meta-analysis, and only Miller (2017) provided appropriate data for the harvest plot, so these studies had a limited effect on the quantitative findings. Furthermore, the inclusion of pretest-posttest studies is not necessarily cause for alarm given that there are so few of them, and only Curran (2018) was assumed to have confounding issues. Nevertheless, noting their presence in this corpus of evidence is important as a possible hinderance in the quality of the evidence.

5.5 Limitations of the review

Though the evidence suggests that school-based interventions which foster academic resilience may reduce inequity and may be effective in improving academic performance, these results should be considered in reference to several limitations. Firstly, this review includes non-randomised studies, which lack the same level of causal inference as experimental designs (Shadish, Cook, & Campbell, 2002). Although this limitation was mitigated by specifying certain types of quasi-experimental designs which may limit confounding, like propensity scoring matching (i.e. Dougherty, 2017), the possibility still remains that causality could have been affected as a result of their inclusion, especially

when considering weaker quasi-experimental designs like pretest-posttest studies. Secondly, as mentioned above, the review relies heavily on the use of evidence that is of mixed quality. Considering the large proportions of both unclear-risk and high-risk studies in the results, study quality serves as an additional threat to validity.

Thirdly, poor reporting practices is a notable limitation. All but one study (Tak, 2014) had no published online protocols. Without transparent reporting standards, it is very possible that researchers can selectively choose to report data that positively support their hypotheses, which introduces bias into the research evidence. This is especially concerning when considering equity evidence. Though most of the authors included in the harvest plots stated that they pre-specified their subgroup analyses, without access to their protocols, there is no clear way of checking that they did indeed specify these analyses *a priori*. There is a possibility that these moderator analyses were performed *a posteriori* and were reported because they happened to produce positive results. Given the stringent criteria that need to met before performing subgroup analyses, including pre-specification, it is vital that authors clearly report any planned analyses in a protocol before they conduct their studies.

Fourthly, educational inequity was only examined through the equity lens of gender, race, and socioeconomic status instead of all eleven PROGRESS-Plus factors. The nature of inequity in education is very dynamic, and perhaps attributing this phenomenon to the 3 most important factors within the scope of this review may oversimplify the nature of the problem. Finally, financial, time, and practical constraints have limited the review author's abilities in conducting this review. Though there was a second reviewer involved in the selection of studies (PD), the data collection, analysis, and writing were all completed by the review author. Without the presence of additional research staff to counterbalance researcher judgement, the review could have been unknowingly influenced by the review

author. Given these limitations, any inferences based on the findings for both effectiveness and equity should be made with caution.

5.6 Strengths of the review

This review had a number of strengths. Firstly, the review presents an innovative conceptual framework that reflects educational sociology, educational psychology, positive psychology, and aspects of epidemiology to explain the problem theory of academic failure and its intercorrelations with educational inequity. Secondly, the review showcases several thoughtful considerations towards ensuring methodological precision. Specifically, the search strategy was highly comprehensive as it was fully informed by evidence-based research on academic protective factors. Cohen's kappa improved the construct validity of the selection criteria which ensured that the correct studies were being captured in the screening process (Westen & Rosenthal, 2005). Additionally, construct validity was further strengthened through the *a priori* development of a data extraction tool which ensured that the right types of data were collected for this review. Given the conceptual complexity of academic resilience, and its overlap with other realms of social intervention like social and emotional learning, there is an increased chance of including incorrect data. Thus, certifying that the data collection methods were highly accurate through the aforementioned steps is a major attribute for this review.

Thirdly, sophisticated analytical methods were used to improve the accuracy of the findings. Effective sample sizes were calculated using an intraclass correlation coefficient for all clustered studies to account for data dependency issues. This reduced the likelihood of a type I error when synthesising effect sizes in the meta-analysis, thus minimising the over-inflation of intervention effectiveness. This review also utilised harvest plots and several robustness checks which are not commonly reported in most reviews. The

robustness checks gave the findings credibility by pinpointing which factors may be compromising the validity of the results, and thus the conclusions gleaned from them. The decision to use harvest plots alongside forest plots is an important aspect of this review. It is equally important to know if interventions are indeed effective and if these effects are as impactful for students who are at greater academic risk.

6 Conclusions

6.1 Research implications

Effectiveness. The work in this review has major implications for how further research could improve the field of academic resilience. More research needs to be done on interventions that have been developed specifically within the conceptual framework of academic resilience. This will directly improve the theories underscoring this research as many included studies did not specify a clear theory of change. Additional research is also needed on all protective factors but especially adaptive help-seeking and self-care behaviours as no study in this review targeted these attributes. Moreover, more research is needed not just in terms of quantity, but also quality given the proportion of unclear- and high-risk studies included in this review so that better evidence can lead to stronger practice and policy decisions. Future studies should evaluate if the effects of these dispositional protective factors can improve academic performance. Though a single protective factor can be used in response to a number of academic risk factors, and vice versa, it would be immensely useful to know if fostering certain psychosocial skills are the most effective means of offsetting particular risks. Using research to determine if the best way to combat academic failure due to poverty is through school-based social support, for example, could have huge implications for decisions on best practice. Given the complexity of attributing a

certain protective factor to a specific academic outcome, future research could utilise factorial study designs and path analyses as ways of establishing better causal inferences between academic protective factors and any subsequent reduction in academic risk. Resilience, even a context-specific form like academic resilience, is a very intricate concept so improving theory generation as well as the amount of evidence-based research can contribute to a standardised way of understanding, evaluating, and using academic resilience programmes in schools.

Equity. This review has highlighted the particular significance, and current shortcomings, of equity in school-based academic resilience interventions. Using research to further the theoretical framework of academic resilience is equally important in terms of equity. Additional research contributions to academic resilience theory could better reflect that academic failure can stem from risk factors like race, gender, and socioeconomic status, which cause systemic and unfair differences in educational attainment. Developing the theory base could provide the conceptual basis needed for better measurement of disadvantage, which could then advance equity assessment methods in school-based research. An example of an applicable conceptual framework is the equity-effectiveness loop (Welch et al., 2008). This model outline 6 steps to assess effectiveness and equity together. Step 1 involves understanding the social determinants of inequity. Step 2 requires outlining community effectiveness strategies by equity factor(s), and how effectiveness may be mediated by forces such as barriers to service and fidelity. The third step considers any potential trade-offs between equity and efficiency based on economic resources while step 4 encompasses knowledge translation processes to put evidence into practice. The final steps, 5 and 6, involve program monitoring and re-assessment. Though this model is public-health focused, it serves as a solid example of the types of conceptual frameworks needed in academic resilience, especially in regards to translating research into practice.

Having better theoretical frameworks could be vital for analytical tools like moderator analyses. Given the stringent requirements that need to be specified *a priori*, having a stronger programme theory can better dictate research decisions on moderator analyses, including the necessary recruitment efforts for sufficient sample sizes to ensure statistical rigour.

Future research specifically needs to address the lack of equity evidence on socioeconomic status. Socioeconomic status and catchment areas are highly correlated, so children in resource-poor communities tend to attend underfunded schools, which widen the educational inequity gap between them and their high-income peers (Aikens & Barbarin, 2008). Therefore, research efforts are needed to determine if academic resilience is effective among low-income students as a way of offsetting any deleterious effects from their surroundings. Finally, this review has significant implications about reporting standards, especially for equity data. If the considerations of equity research are to inform practice and policy decisions, then the evidence must find its way into publications. This requires more transparent and standardised methods of reporting. More protocols need to be published before conducting studies so that the reported outcomes in the final studies can be held to a higher level of rigour. The development of a CONSORT-equity guideline, akin to the reporting standards from Welch and colleagues (2015), is needed for non-health related equity research such as this review. Moreover, given the difficulties associated with conducting RCTs in school settings, such as ethical concerns and school policies, an equity reporting guideline must be created for non-randomised trials as well (Baron et al., 2017). Taken together, these implications and recommendations can assist in making equity as a prominent feature in future social intervention research. Though effectiveness is indeed important, it only matters insofar as the researchers' ability to determine *'for whom'* these interventions work.

6.2 Practice implications

This review has provided valuable insights into how to effectively use these school-based programmes to address both academic underperformance and educational inequity. Implications about practice for effectiveness and equity will be discussed in tandem. Considering the multiple ways to improve educational performance through the use of academic resilience, the wide array of interventions presented in this review can be used to cater to students of differing backgrounds and needs. The variety of interventions available can be particularly useful to school practitioners. School psychologists who work closely with students from many different demographics are in a unique position to institute interventions that can strengthen students' academic capabilities (Moolla & Lazarus, 2014). In this way, academic resilience interventions can be both effective and equitable at bolstering school success given the number of tools at a practitioner's disposal.

Furthermore, academic resilience interventions are flexible – since they use psychosocial tools as opposed to educational tools, they are not tethered to certain academic content. They can be used in a myriad of academic courses, settings, and school levels. Academic resilience interventions can be modified within reason to fit differing educational needs and objectives which allow school practitioners such as teachers, principals, and school psychologists to use one intervention in multiple ways. This is particularly useful for addressing equity needs in underprivileged schools where it is not feasible to implement several programmes at once given financial and resource constraints.

Practice decisions should also consider when to intervene in a student's academic career. The middle school years, approximately 11 to 13 years of age, is the most represented student population in this corpus of evidence. Thus, this phase in a student's life may be the most opportune time for intervention before the transition to secondary school causes further academic downturn due to increased academic difficulty and changes

in social and learning contexts (Hanewald, 2013; Martínez et al., 2011). It is important that intervention efforts do not begin and end in middle school, but instead continue on throughout a student's academic life-course as needed. Tying back to Morales' (2000) Resilience Cycle from the introduction (section 1.4), constant use of these interventions sustains students' cognitive-behavioural patterns supporting their newfound academic achievement until their desired goal has been reached. Once their academic goal has been realised, for example progressing to a new grade level, students can then refine the skills they have learned through these programmes to overcome future academic hurdles.

In addition to deciding on the right time to intervene, best practice judgments should include implementation strategies. Interactive delivery methods were used in many of the included studies which involved active participation between the program leader, typically the teacher, and the students. This created a positive learning environment whereby students not only acquired protective factors, they shared a communal educational experience which in itself may have increased academic abilities. Especially for students who come from impoverished backgrounds and have a poor home learning environment, it is important that they receive as much as possible from these school-based interventions. Thus, where possible, these interventions could be delivered interactively to elicit a stronger impact. This delivery strategy, however, requires proper implementation training for practitioners such as teachers and support staff. Schools that are committed to improving students' academic performance must ensure that their personnel receive the right support in implementing these programmes. Considering the equity-effectiveness loop once more, the importance of knowledge translation methods, including implementation efforts, cannot be overlooked.

6.3 Policy implications

The effectiveness and equity of academic resilience interventions must be considered within the context of evidence-based policy. Though additional research could better inform both effectiveness and equity impacts, as the current evidence stands, policymakers should begin to develop policy frameworks to support academic resilience interventions as they have demonstrated positive effects on academic gains. Additionally, the review has shown some potential in reducing inequities for minority students and female students in regards to specific courses like maths. Again, it is important to stress that additional evidence is needed in regards to equity impacts, especially for socioeconomic status. However, using the limited equity evidence presented in this review could highlight vital evidence gaps which can lead to policy agendas that encourage better research which in turn informs better practice.

The effective and equitable impact of these interventions are particularly noteworthy in light of mounting school pressure from policymakers to increase students' academic performance using evidence-based research (DiGiacomo, 2014). However, to maintain academic gains, additional school reform is needed which embeds academic resilience programmes into the school curriculum. This requires policy frameworks that have an infrastructure for these school-based programmes to become visible, sustained, and recognisable (Ross et al., 2001). Drawing on existing policy, these interventions could fit well into current education legislative agendas for response to intervention in the United States (U.S. Department of Education, 2003). Response to intervention is a multi-tiered educational instruction approach defined by a system of academic and psychosocial support tailored to student's needs. (Sailor, 2009). Additionally, the No Child Left Behind Act of 2001 requires the use of research-based education to close performance gaps so students perform at state and federal standards (U.S. Department of Education, 2011).

Given the range of academic resilience programmes, flexibility of use, and various prevention levels, they could easily be integrated into these existing educational policy approaches.

In terms of practicality, if academic resilience interventions are added to current or future policies, attention needs to be paid to cost-effectiveness. Cost-benefit analyses comparing the impact of these interventions versus tradition educational programmes should inform policymaking decisions surrounding budgets and resource allocation. Since academic resilience is an emerging research field, no policy currently exists on resource considerations, thus creating a policy gap that can be addressed in future research. Regarding educational inequity, this will be of particular importance in resource-poor contexts, which are constantly underperforming because they receive less federal funding. Looking at the American education system, funding decisions are typically dictated by performance indicators, such as test scores and national rankings (U.S. Department of Education, 2011). This results in a vicious circle of underperformance and underfunding. Therefore, if policy is to shift the current educational landscape by levelling the playing field, there needs to be an emphasis on providing the structure and resources for poor schools to effectively use these interventions.

Academic resilience interventions could also provide insights for informing educational policy on an international scale. Given the global purview of this research, its findings have major implications for exacting international education policy points. Akin to the U.S. No Child Left Behind Act, the U.N. Sustainable Development Goal 4 aims to close global achievement gaps by promoting inclusive, quality education for all, especially for women and girls. Though academic resilience cannot address the ongoing accessibility issues to quality education around the world, once students are enrolled in school, it can be used as a worthwhile tool to sustain their learning. Academic resilience interventions could

be potentially useful in eliminating gender disparities in education as this review has showcased the gravity of academic risk by gender and how these programmes can reduce gender-based academic failure through the use of academic psychosocial protective factors (United Nations, 2016). Connecting back to the theories of gender inequity outlined in the introduction, female students in developing countries are less likely to receive an education relative to their male counterparts. Thus, academic resilience could help to reduce the gender gap in education by lessening the impact of the academic risk factors that girls face in these low-income contexts. Thus, academic resilience has both contributed to existing policy frameworks but also offers innovative policy recommendations for tackling educational underperformance and inequity on a global scale.

7 Appendices

7.1 Table of excluded studies

Table of excluded studies: https://bit.ly/2ImHc8t

7.2 Data extraction form

Data extraction form: http://bit.ly/2Ml96E8

7.3 Equity data matrix

Equity data matrix: http://bit.ly/2lHY0xP

7.4 RoB 2.0

RoB 2.0 sample form: http://bit.ly/2Kx8URy

7.5 ROBINS-I

ROBINS-I sample form: http://bit.ly/2lJmTco

7.6 Sample R script

All R script: http://bit.ly/2KgLrbb

7.7 Authors' correspondence list

Correspondence list: http://bit.ly/2tzZgr5

7.8 Ethics

Approval letter: http://bit.ly/2lENIOR

8 References

8.1 References of included studies

Arif, M. I., & Mirza, M. (2017). Effectiveness of an Intervention Program in Fostering Academic Resilience of Students at Risk of Failure at Secondary School Level. *Bulletin of Education & Research*, *39*(1), 251–264.

Asmus, J. M., Carter, E. W., Moss, C. K., Biggs, E. E., Bolt, D. M., Born, T. L., … L.B., V. (2017). Efficacy and Social Validity of Peer Network Interventions for High School Students With Severe Disabilities. *American Journal on Intellectual and Developmental Disabilities*, *122*(2), 118–137.

Brigman, G., & Campbell, C. (2003). Helping Students Improve Academic Achievement and School Success Behavior. *Professional School Counseling*, *7*(2), 91–98.

Cohen, G. L. (2006). Reducing the Racial Achievement Gap: A Social-Psychological Intervention. *Science*, *313*(5791), 1307–1310.

Curran, S. (2018). The effects of flexible ability grouping on mathematics improvement and self-concepts in the intermediate grades. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, *78*(11–A(E)).

DiGiacomo, G. (2014). Enhancing self-monitoring and self-reflection through a self-regulatory skills intervention embedded in a middle school mathematics curriculum. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, *75*(10–A(E)).

Dupuis, K. L. (2013). Impacts of a school-based mentoring program on academic and behavioral outcomes for middle school youth. *ProQuest Dissertations and Theses*, *74*(1–A(E)).

Dougherty, D., & Sharkey, D. (2017). Reconnecting Youth: Promoting emotional competence and social support to improve academic achievement. *Children and Youth Services Review*, *74*, 28–34.

Eggert, L. L., & Seyl, C. D. (1990). Effects of a school-based prevention program for potential high school dropouts and drug abusers. *International Journal of the Addictions*, *25*(7), 773–801.

Eskreis-Winkler, L., Shulman, E. P., Young, V., Tsukayama, E., Brunwasser, S. M., & Duckworth, A. L. (2016). Using wise interventions to motivate deliberate practice. *Journal of Personality and Social Psychology*, *111*(5), 728–744.

Felner, R. D., Ginter, M., & Primavera, J. (1982). Primary prevention during school transitions: social support and environmental structure. *American Journal of Community Psychology*, *10*(3), 277–290.

Harrison, S. E., Li, X., Zhang, J., Chi, P., Zhao, J., & Zhao, G. (2017). Improving school

outcomes for children affected by parental HIV/AIDS: Evaluation of the ChildCARE Intervention at 6-, 12-, and 18-months. *School Psychology International*, *38*(3), 264–286.

Haynes, N. M. (1990). The effects of a school development program on self-concept. *Yale Journal of Biology and Medicine*, *63*(4), 275–283.

Hu, W., Adey, P., Jia, X., Liu, J., Zhang, L., Li, J., & Dong, X. (2011). Effects of a 'Learn to Think' intervention programme on primary school students. *The British Journal of Educational Psychology*, *81*(Pt 4), 531–557.

Karcher, M. J. (2008). The study of mentoring in the learning environment (SMILE): a randomized evaluation of the effectiveness of school-based mentoring. *Prevention Science: The Official Journal of the Society for Prevention Research*, *9*(2), 99–113.

Miller, C. F., Kochel, K. P., Wheeler, L. A., Updegraff, K. A., Fabes, R. A., Martin, C. L., … C.L., M. (2017). The efficacy of a relationship building intervention in 5th grade. *Journal of School Psychology*, *61*, 75–88.

Muratori, P., Bertacchi, I., Giuli, C., Nocentini, A., Ruglioni, L., & Lochman, J. E. (2016). Coping Power Adapted as Universal Prevention Program: Mid Term Effects on Children's Behavioral Difficulties and Academic Grades. *The Journal of Primary Prevention*, *37*(4), 389–401.

Paunesku, D., Walton, G. M., Romero, C., Smith, E. N., Yeager, D. S., & Dweck, C. S. (2015). Mind-set interventions are a scalable treatment for academic underachievement. *Psychological Science*, *26*(6), 784–793.

Portwood, S. G., Ayers, P. M., Kinnison, K. E., Waris, R. G., Wise, D. L., S.G., P., … R.G., W. (2005). YouthFriends: Outcomes from a school-based mentoring program. *Journal of Primary Prevention*, *26*(2), 129–145.

Reyes, O., & Gillock, K. (1994). A longitudinal study of school adjustment in urban, minority adolescents: effects of a high school transition program. *American Journal of Community Psychology*, *22*(3), 341–369.

Shoshani, A., Steinmetz, S., & Kanat-Maymon, Y. (2016). Effects of the Maytiv positive psychology school program on early adolescents' well-being, engagement, and achievement. *Journal of School Psychology*, *57*, 73–92.

Snyder, F., Flay, B., Vuchinich, S., Acock, A., Washburn, I., Beets, M., & Li, K.-K. (2009). Impact of a social-emotional and character development program on school-level indicators of academic achievement, absenteeism, and disciplinary outcomes: A matched-pair, cluster randomized, controlled trial. *Journal of Research on Educational Effectiveness*, *3*(1), 26–55.

Tak, Y. R., Kleinjan, M., Lichtwarck-Aschoff, A., & Engels, R. C. M. E. (2014). Secondary outcomes of a school-based universal resiliency training for adolescents: A cluster randomized controlled trial. *BMC Public Health*, *14*(1).

Terry, J. (2016). *Preliminary evaluation of "Footprints:" Motivational interviewing to promote cognitive-behavioral skills, academic outcomes, and academic protective factors in middle school students. ProQuest Dissertations and Theses*. University of South Carolina, Ann Arbor.

8.2 Supplements

Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Supporting Online Material for Reducing the racial achievement gap: A social-psychological intervention. *Science*, *313*(5791), 1307–1310.

Eskreis-Winkler, L. (2016). Online Supplement. Retrieved from http://supp.apa.org/psycarticles/supplemental/pspp0000074/pspp0000074_supp.html

Miranda, A., Brigman, G., & Peluso, P. (2007). Student Success Skills : A Promising Program to Close the Academic Achievement Gap, (2003), 490–497.

Tak, Y. R., Van Zundert, R. M., Kuijpers, R. C., Van Vlokhoven, B. S., Rensink, H. F., & Engels, R. C. (2012). A randomized controlled trial testing the effectiveness of a universal school-based depression prevention program 'Op Volle Kracht' in the Netherlands. *BMC Public Health*, *12*, 21.

8.3 Additional references

Aikens, N. L., & Barbarin, O. (2008). Socioeconomic differences in reading trajectories: The contribution of family, neighborhood, and school contexts. *Journal of Educational Psychology*, *100*(2), 235–251. https://doi.org/10.1037/0022-0663.100.2.235

Alva, S. A. (1991). Academic Invulnerability Among Mexican-American Students: The Importance of Protective Resources and Appraisals. *Hispanic Journal of Behavioral Sciences*, *13*(1), 18–34. https://doi.org/10.1177/07399863910131002

Ambady, N., Shih, M., Kim, A., & Pittinsky, T. L. (2001). Stereotype susceptibility in children: Effects of Identity Activation on Quantitative Performance. *Psychological Science*, *12*(5), 385–390. https://doi.org/10.1111/1467-9280.00371

Appleyard, K., Egeland, B., van Dulmen, M. H. M., & Sroufe, L. A. (2005). When more is not better: The role of cumulative risk in child behavior outcomes. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *46*(3), 235–245. https://doi.org/10.1111/j.1469-7610.2004.00351.x

Arif, M. I., & Mirza, M. (2017). Effectiveness of an Intervention Program in Fostering Academic Resilience of Students at Risk of Failure at Secondary School Level. *Bulletin of Education & Research*, *39*(1), 251–264. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eax&AN=125765141&site=ehost-live

Austin, B. W. (1996). *Repairing the break: Key ways to support family life, reclaim our streets, and rebuild civil society in America's communities.*

Baldwin, S. A., Murray, D. M., & Shadish, W. R. (2005). Empirically supported treatments or type I errors? Problems with the analysis of data from group-administered treatments. *Journal of Consulting and Clinical Psychology*, *73*(5), 924–935. https://doi.org/10.1037/0022-006X.73.5.924

Bandura, A. (1997). Self-efficacy: The exercise of control. New York: W.H. Freeman and Company. *American Psychological Association*, *23*, 604.

Bandura, A., Freeman, A., Mahoney, M. J., DeVito, P., & Martin, D. (2004). Model of causality in social learning theory. In *Cognition and psychotherapy (2nd ed.)*. (pp. 25–44). https://doi.org/10.1007/978-1-4684-7562-3_3

Banerjee, P. A. (2017). Does continued participation in STEM enrichment and enhancement activities affect school maths attainment? *Oxford Review of Education*, *43*(1), 1–18. https://doi.org/10.1080/03054985.2016.1235031

Barham, V., Boadway, R., Marchand, M., & Pestieau, P. (1995). Education and the poverty trap. *European Economic Review*, *39*(7), 1257–1275. https://doi.org/10.1016/0014-2921(94)00040-7

Baron, A., Evangelou, M., Malmberg, L.-E., & Melendez-Torres, G. J. (2017). The Tools of the Mind curriculum for improving self-regulation in early childhood: a sytematic review, 1–77.

Beilock, S. L. (2008). Math performance in stressful situations. *Current Directions in Psychological Science*. https://doi.org/10.1111/j.1467-8721.2008.00602.x

Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences*, *107*(5), 1860–1863. https://doi.org/10.1073/pnas.0910967107

Belsky, J., Bakermans-Kranenburg, M. J., & Van Ijzendoorn, M. H. (2007). For better and for worse: Differential susceptibility to environmental influences. *Current Directions in Psychological Science*, *16*(6), 300–304. https://doi.org/10.1111/j.1467-8721.2007.00525.x

Biafora, F. A., Warheit, G. J., Zimmerman, R. S., Gil, A. G., Apospori, E., Taylor, D., & Vega, W. A. (1993). Racial Mistrust and Deviant Behaviors Among Ethnically Diverse Black Adolescent Boys. *Journal of Applied Social Psychology*, *23*(11), 891–910. https://doi.org/10.1111/j.1559-1816.1993.tb01012.x

Blake, J. (1989). *Family size and achievement. Berkeley*. University of California Press.

Borman, G. D., & Overman, L. T. (2004). Academic Resilience in Mathematics among Poor and Minority Students. *The Elementary School Journal*, *104*(3), 177–195. https://doi.org/10.1086/499748

Bowles, S., & Gintis, H. (2002). The inheritance of inequality. *Journal of Economic Perspectives*, *16*(3), 3–30. https://doi.org/10.1257/089533003769204434

Bradley, K. (2014). Gender Inequality in Education. In *Education and sociology: An encyclopedia*.

Brigman, G., & Campbell, C. (2003). Helping Students Improve Academic Achievement and School Success Behavior. *Professional School Counseling*, *7*(2), 91–98. Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc4&NEWS=N&AN=2004-10595-004

Bronfenbrenner, U. (1979). Ecological systems theory. *Annals of Child Development*, *6*, 187–249. Retrieved from http://psycnet.apa.org/psycinfo/1992-98662-005

Brooks-Gunn, J., Duncan, G., & Aber, J. L. (1997). *Neighborhood poverty: Context and consequences for development*. New York: Russell Sage Foundation.

Brown, J. L., Jones, S. M., LaRusso, M. D., & Aber, J. L. (2010). Improving Classroom Quality: Teacher Influences and Experimental Impacts of the 4Rs Program. *Journal of Educational Psychology*, *102*(1), 153–167. https://doi.org/10.1037/a0018160

Brown, R. P., & Pinel, E. C. (2003). Stigma on my mind: Individual differences in the experience of stereotype threat. *Journal of Experimental Social Psychology*, *39*(6), 626–633. https://doi.org/10.1016/S0022-1031(03)00039-8

Campbell and Cochrane Equity Methods Group. (n.d.). Our Publications. Retrieved November 30, 2017, from http://methods.cochrane.org/equity/our-publications

Cappella, E., & Weinstein, R. S. (2001). Turning around reading achievement: Predictors of high school students' academic resilience. *Journal of Educational Psychology*, *93*(4), 758–771. https://doi.org/10.1037/0022-0663.93.4.758

CASEL. (n.d.). What is SEL? Retrieved April 28, 2018, from https://casel.org/what-is-sel/

CASEL. (2012). *Effective Social and Emotional Learning Programs. Preschool and Elementary School Edition*. https://doi.org/http://casel.org/wp-content/uploads/2016/01/2013-casel-guide-1.pdf

Caspi, A., Wright, B. R. E., Moffitt, T. E., & Silva, P. A. (1998). Early Failure in the Labor Market: Childhood and Adolescent Predictors of Unemployment in the Transition to Adulthood. *American Sociological Review*, *63*(3), 424. https://doi.org/10.2307/2657557

Cassidy, S. (2016). The Academic Resilience Scale (ARS-30): A new multidimensional construct measure. *Frontiers in Psychology*, *7*(NOV). https://doi.org/10.3389/fpsyg.2016.01787

Cicchetti, D., & Curtis, W. J. (2007). Multilevel perspectives on pathways to resilient functioning. *Development and Psychopathology*, *19*(03), 627. https://doi.org/10.1017/S0954579407000314

Coard, B. (1971). How the West Indian Child is Made Educationally Subnormal in the British School System: The Scandal of the Black Child in Schools in Britain.

Cohen, G. L. (2006). Reducing the Racial Achievement Gap: A Social-Psychological Intervention. *Science*, *313*(5791), 1307–1310. https://doi.org/10.1126/science.1128317

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. *Statistical Power Analysis for the Behavioral Sciences*. https://doi.org/10.1234/12345678

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Colp, S. M., & Nordstokke, D. W. (2014). Exploring the measurement of academic resilience. In *Proceedings of the Symposium presentation at the Canadian Psychological Association's 75th Annual Convention*.

Cook, D. (2014). The Gender Gap in Math Is Not Universal. Retrieved from

http://chance.amstat.org/2014/11/visiphilia/

Crosnoe, R., Mistry, R. S., & Elder, G. H. J. (2002). Economic Disadvantage , Family Dynamics , in Higher Adolescent Enrollment Education. *Journal of Marriage and Family*, *64*(3), 690–702. https://doi.org/10.1111/j.1741-3737.2002.00690.x

Curran, S. (2018). The effects of flexible ability grouping on mathematics improvement and self-concepts in the intermediate grades. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, *78*(11–A(E)), No-Specified. Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc14&NEWS=N&AN=2017-36662-062

Davis-Kean, P. E. (2005). The Influence of Parent Education and Family Income on Child Achievement: The Indirect Role of Parental Expectations and the Home Environment. *Journal of Family Psychology*, *19*(2), 294–304. https://doi.org/10.1037/0893-3200.19.2.294

Deeks, J. J. (2011). *Chapter 9: Analysing data and undertaking. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011).*

DHS. (2003). Statcompiler, ORC Macro. Measure DHS+ project. Retrieved from https://www.statcompiler.com/en/

DiGiacomo, G. (2014). Enhancing self-monitoring and self-reflection through a self-regulatory skills intervention embedded in a middle school mathematics curriculum. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, *75*(10–A(E)), No-Specified. Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc12&NEWS=N&AN=2015-99071-145

Dishion, T. J., & Connell, A. (2006). Adolescents' resilience as a self-regulatory process: Promising themes for linking intervention with developmental science. In *Annals of the New York Academy of Sciences* (Vol. 1094, pp. 125–138). https://doi.org/10.1196/annals.1376.012

Duncan, T. G., & McKeachie, W. J. (2005). The Making of the Motivated Strategies for Learning Questionnaire. *Educational Psychologist*, *40*(2), 117–128. https://doi.org/10.1207/s15326985ep4002_6

Durlak, J. A. (1998). Why Program Implementation is Important. *Journal of Prevention & Intervention in the Community*, *17*(2), 5–18. https://doi.org/10.1300/J005v17n02_02

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, *41*(3–4), 327–350. https://doi.org/10.1007/s10464-008-9165-0

Eckenrode, J., Laird, M., & Doris, J. (1993). School performance and disciplinary problems among abused and neglected children. *Developmental Psychology*, *29*(1), 53–62. https://doi.org/10.1037/0012-1649.29.1.53

Egger, M., Davey Smith, G., Schneider, M., Minder, C., Mulrow, C., Egger, M., … Olkin, I. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal (Clinical Research Ed.)*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

Farver, J. A. M., Xu, Y., Eppe, S., & Lonigan, C. J. (2006). Home environments and young Latino children's school readiness. *Early Childhood Research Quarterly*, *21*(2), 196–212. https://doi.org/10.1016/j.ecresq.2006.04.008

Felliti, V., & Anda, R. (1997). The Adverse Childhood Experiences (ACE) study. *Atlanta, GA: Center for Disease Control and Prevention …*, 1–2.

Ferguson, H. B., Bovaird, S., & Mueller, M. P. (2007). The impact of poverty on educational outcomes for children. *Paediatrics and Child Health*, *12*(8), 701–706.

Ferguson, R. F. (2003). Teachers' Perceptions and Expectations and the Black-White Test Score Gap. *Urban Education*, *38*(4), 460–507. https://doi.org/10.1177/0042085903038004006

Fishman, J. (1997). Language and Ethnicity: The View from Within. In *The Handbook of Sociolinguistics* (pp. 327–343).

Flay, B. R., & Allred, C. G. (2010). The Positive Action Program: Improving Academics, Behavior, and Character by Teaching Comprehensive Skills for Successful Learning and Living. In *International Research Handbook on Values Education and Student Wellbeing* (pp. 471–501). https://doi.org/10.1007/978-90-481-8675-4_28

Ford, E. S. (2011). Impact of a school-based behavioral modification program on indicators of academic achievement for secondary students with serious emotional disturbance. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, *72*(5–A), 1613. Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc8&NEWS=N&AN=2011-99210-173

Garmezy. (1991). Resiliency and Vulnerability to Adverse Developmental Outcomes Associated with Poverty. *The American Behavioral Scientist*, *34*(4), 416–430. https://doi.org/10.1017/CBO9781107415324.004

Gordon, E. W., & Song, L. D. (1994). Variations in the experience of resilience. In *Educational Resilience in Inner-city America: Challenges and Prospects* (pp. 27–43).

Gordon, K. (1995). Self-Concept and Motivational Patterns of Resilient African American High School Students. *Journal of Black Psychology*, *21*(3), 239–255. https://doi.org/10.1177/00957984950213003

Gordon, K. (1996). Resilient Hispanic Youths' Self-Concept and Motivational Patterns. *Hispanic Journal of Behavioral Sciences*, *18*(1), 63–73. https://doi.org/10.1177/07399863960181007

Grantham, T. C. (2004). Multicultural Mentoring to Increase Black Male Representation in Gifted Programs. *Gifted Child Quarterly*, *48*(3), 232–245. https://doi.org/10.1177/001698620404800307

Green, L. W., & Glasgow, R. E. (2006). Evaluating the relevance, generalization, and applicability of research: Issues in external validation and translation methodology. *Evaluation and the Health Professions*. https://doi.org/10.1177/0163278705284445

Guyll, M., Madon, S., Prieto, L., & Scherr, K. C. (2010). The potential roles of self-fulfilling prophecies, stigma consciousness, and stereotype threat in linking Latino/a ethnicity and educational outcomes. *Journal of Social Issues*, *66*(1), 113–130.

https://doi.org/10.1111/j.1540-4560.2009.01636.x

Gwatkin, D. R. (2003). How well do health programmes reach the poor? *Lancet*. https://doi.org/10.1016/S0140-6736(03)12558-5

Hanewald, R. (2013). Transition Between Primary and Secondary School: Why it is Important and How it can be Supported. *Australian Journal of Teacher Education*, *38*(1). https://doi.org/10.14221/ajte.2013v38n1.7

Harrison, S. E., Li, X., Zhang, J., Chi, P., Zhao, J., & Zhao, G. (2017). Improving school outcomes for children affected by parental HIV/AIDS: Evaluation of the ChildCARE Intervention at 6-, 12-, and 18-months. *School Psychology International*, *38*(3), 264–286. https://doi.org/http://dx.doi.org/10.1177/0143034316689589

Hartley, M. T. (2011). Examining the relationships between resilience, mental health, and academic persistence in undergraduate college students. *Journal of American College Health*, *59*(7), 596–604. https://doi.org/10.1080/07448481.2010.515632

Hawe, P., Shiell, A., & Riley, T. (2004). Complex interventions: how "out of control" can a randomised controlled trial be? *BMJ : British Medical Journal*, *328*(7455), 1561–1563. https://doi.org/10.1136/bmj.328.7455.1561

Hedges, L. V., & Hedberg, E. C. (2007). Intraclass Correlation Values for Planning Group-Randomized Trials in Education. *Educational Evaluation and Policy Analysis*, *29*(1), 60–87. https://doi.org/10.3102/0162373707299706

Higgins, J., Altman, D., & Sterne, J. (Editors). (2011). Chapter 8: Assessing risk of bias in included studies. In *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011)*. (p. Available from www.cochrane-handbook.org). https://doi.org/10.1001/jama.2013.109501.Conflict

Higgins, J., Sterne, J., Savović, J., Page, M., Hróbjartsson, A., Boutron, I., … Eldridge, S. (2016). A revised tool for assessing risk of bias in randomized trials. In *Cochrane Database of Systematic Reviews* (p. Issue 10 (Suppl 1)).

Hillsdon, M., Foster, C., Naidoo, B., & Crombie, H. (2004). The effectiveness of public health interventions for increasing physical activity among adults : A review of reviews. *Evidence Briefing*. Retrieved from www.hda.nhs.uk

Hirsch, D. and J. A. (2007). Experiences of poverty and educational disadvantage. *Education*, (32), 10. https://doi.org/10.1787/715165322333

Hrabowski III, A. I. I. I., Maton, I., & Greif, L. (1998). Beating the odds: Raising academically successful African American males. *Beating the Odds: Raising Academically Successful African American Males*.

Hu, W., Adey, P., Jia, X., Liu, J., Zhang, L., Li, J., & Dong, X. (2011). Effects of a "Learn to Think" intervention programme on primary school students. *The British Journal of Educational Psychology*, *81*(Pt 4), 531–557. https://doi.org/http://dx.doi.org/10.1348/2044-8279.002007

Hudley, C., & Graham, S. (2001). Stereotypes of Achievement Striving Among Early Adolescents. *Social Psychology of Education*, *5*(2), 201–224. https://doi.org/10.1023/A:1014438702266

Humphreys, D. K., & Ogilvie, D. (2013). Synthesising evidence for equity impacts of population-based physical activity interventions: a pilot study. *International Journal of Behavioral Nutrition and Physical Activity*, *10*(1), 76. https://doi.org/10.1186/1479-5868-10-76

Irving, M. A., & Hudley, C. (2005). Cultural mistrust, academic outcome expectations, and outcome values among African American adolescent men. *Urban Education*, *40*(5), 476–496.

Jackson, J. F. L., & Moore, J. L. (2006). African American males in education: Endangered or ignored? *Teachers College Record*, *108*(2), 201–205. https://doi.org/10.1111/j.1467-9620.2006.00647.x

Jacob, J. C. (1981). Theories of Social and Educational Inequality: From dichotomy to typology. *British Journal of Sociology of Education*, *2*(1), 71–89. https://doi.org/10.1080/0142569810020105

Jenkins, M. (2015). Too Many Kids. *The Atlantic*.

Johnson, M. L., Taasoobshirazi, G., Kestler, J. L., & Cordova, J. R. (2015). Models and messengers of resilience: a theoretical model of college students' resilience, regulatory strategy use, and academic achievement. *Educational Psychology*, *35*(7), 869–885. https://doi.org/10.1080/01443410.2014.893560

Jowkar, B., Kojuri, J., Kohoulat, N., & Hayat, A. A. (2014). Academic resilience in education: the role of achievement goal orientations. *Journal of Advances in Medical Education & Professionalism*, *2*(1), 33–38.

Kahn, E., Ramsey, L., Brownson, R., Heath, G., Howze, E., Powell, K., … Corso, P. (2002). The effectiveness of interventions to increase physical activity: A systematic review. *American Journal of Preventive Medicine*, *22*(4), 73–107. https://doi.org/10.1016/S0749-3797(02)00434-8

Klepinger, D. H., Lundberg, S., & Plotnick, R. D. (1995). Adolescent Fertility and the Educational Attainment of Young Women. *Family Planning Perspectives*, *27*(1), 23–28. https://doi.org/10.2307/2135973

Kohlberg, L., LaCrosse, J., & Ricks, D. (1972). The predictability of adult mental health from childhood behavior. In *Manual of child psychopathology* (pp. 1217–1284).

Ladson-Billings, G., & Tate, W. F. (1995). Toward a critical race theory of education. *Teachers College Record*, *97*(1), 47. https://doi.org/10.1080/10282580701850413

Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, *33*(1), 159. https://doi.org/10.2307/2529310

Leary, K. A., & DeRosier, M. E. (2012). Factors Promoting Positive Adaptation and Resilience during the Transition to College. *Psychology*, *03*(12), 1215–1222. https://doi.org/10.4236/psych.2012.312A180

Leiter, J., & Johnsen, M. C. (1994). Child Maltreatment and School Performance. *Source: American Journal of Education American Journal of Education This*, *102*(2), 154–189. https://doi.org/10.2307/1163250

Leon, D. A., Walt, G., & Gilson, L. (2001). Recent Advances: International Perspectives on

Health Inequalities and Policy. *British Medical Journal*, *322*(7286), 591–594. Retrieved from http://www.bmj.com

Levine, A., & Nidiffer, J. (1996). *Beating the Odds: How the Poor Get to College*. San Francisco, CA, US: The Jossey Bass Higher and Adult Education Series.

Liddle, H. A. (1994). Contextualizing resiliency. In *Educational resilience in inner-city America: Challenges and prospects*. (pp. 167–177). Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1994-97364-012&site=ehost-live&scope=site%5Cnhttp://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1994-97364-012&site=ehost-live

Lloyd, C. B., & Blanc, A. K. (1996). Children's Schooling in Sub Saharan Africa: The Role of Fathers, Mothers and Others. *POPULATION AND DEVELOPMENT REVIEW*, *22*(2), 265–298.

Lorenc, T., & Oliver, K. (2014). Adverse effects of public health interventions: a conceptual framework. *Journal of Epidemiology and Community Health*, *68*(3), 288–290. https://doi.org/10.1136/jech-2013-203118

Luthar, S. S. (1991). Vulnerability and Resilience: A Study of High- Risk Adolescents. *Child Development*, *62*(3), 600–616. https://doi.org/10.1111/j.1467-8624.1991.tb01555.x

Luthar, S. S., Cicchetti, D., & Becker, B. (2000). The construct of resilience: A critical evaluation and guidelines for future work. *Child Development*. https://doi.org/10.1111/1467-8624.00164

Lynch, K., & O'riordan, C. (1998). Inequality in Higher Education: a study of class barriers. *British Journal of Sociology of Education*, *19*(4), 445–478. https://doi.org/10.1080/0142569980190401

Magnuson, K. A., Sexton, H. R., Davis-Kean, P. E., & Huston, A. C. (2009). Increases in Maternal Education and Young Children's Language Skills. *Merrill-Palmer Quarterly*, *55*(3), 319–350. https://doi.org/10.1353/mpq.0.0024

Marmot, M., & Health. (2008). Closing the gap in a generation: health equity through action on the social determinants of health. Final report of the Commission on Social Determinants of Health. *Global Health Promotion*, *Suppl 1*(5), 420. Retrieved from http://discovery.ucl.ac.uk/1305797/

Martin, A. J., & Marsh, H. W. (2006). Academic resilience and its psychological and educational correlates: A construct validity approach. *Psychology in the Schools*, *43*(3), 267–281. https://doi.org/10.1002/pits.20149

Martin, A. J., & Marsh, H. W. (2009). Academic resilience and academic buoyancy: Multidimensional and hierarchical conceptual framing of causes, correlates and cognate constructs. *Oxford Review of Education*, *35*(3), 353–370. https://doi.org/10.1080/03054980902934639

Martínez, R. S., Aricak, O. T., Graves, M. N., Peters-Myszak, J., & Nellis, L. (2011). Changes in Perceived Social Support and Socioemotional Adjustment Across the Elementary to Junior High School Transition. *Journal of Youth and Adolescence*, *40*(5), 519–530. https://doi.org/10.1007/s10964-010-9572-z

Masten, A. S. (1994). Resilience in individual development: Successful adaptation despite risk and adversity. In *Educational resilience in inner-city America: Challenges and prospects.* (pp. 3–25). Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc3&NEWS=N&AN=1994-97364-001

Masten, A. S. (2013). *Risk and Resilience in Development* (Vol. 1). https://doi.org/10.1093/oxfordhb/9780199958474.013.0023

Masten, A. S., Best, K. M., & Garmezy, N. (1990). Resilience and development: Contributions from the study of children who overcome adversity. *Development and Psychopathology*, *2*(04), 425. https://doi.org/10.1017/S0954579400005812

Masten, A. S., Long, J. D., Kuo, S. l-C., McCormick, C. M., & Desjardins, C. D. (2009). Developmental models of strategic intervention. *European Journal of Developmental Science*, *3*(3), 282–291. https://doi.org/10.3233/DEV-2009-3306

McClelland, M. M., Acock, A. C., & Morrison, F. J. (2006). The impact of kindergarten learning-related skills on academic trajectories at the end of elementary school. *Early Childhood Research Quarterly*, *21*(4), 471–490. https://doi.org/10.1016/j.ecresq.2006.09.003

Meier, K., Stewart, J., & England, R. (1998). *Race, class and education: The politics of second generation discrimination.* Madison: University of Wisconsin Press.

Miller, C. F., Kochel, K. P., Wheeler, L. A., Updegraff, K. A., Fabes, R. A., Martin, C. L., … C.L., M. (2017). The efficacy of a relationship building intervention in 5th grade. *Journal of School Psychology*, *61*, 75–88. https://doi.org/http://dx.doi.org/10.1016/j.jsp.2017.01.002

Miller, L. S. (1997). *An American Imperative: Accelerating Minority Educational Advancement*.

Milofsky, C. (1974). Why special education isn't special. *Harvard Educational Review*, *44*(4), 437–458.

Miranda, A., Brigman, G., & Peluso, P. (2007). Student Success Skills : A Promising Program to Close the Academic Achievement Gap, (2003), 490–497.

Moeyaert, M., Ugille, M., Natasha Beretvas, S., Ferron, J., Bunuan, R., & Van den Noortgate, W. (2017). Methods for dealing with multiple outcomes in meta-analysis: a comparison between averaging effect sizes, robust variance estimation and multilevel meta-analysis. *International Journal of Social Research Methodology*, *20*(6), 559–572. https://doi.org/10.1080/13645579.2016.1252189

Moolla, N., & Lazarus, S. (2014). School psychologists' views on challenges in facilitating school development through intersectoral collaboration. *South African Journal of Education*, *34*, 1–10.

Morales, E. E. (2000). A contextual understanding of the process of educational resilience: High achieving Dominican American students and the "Resilience Cycle." *Innovative Higher Education*, *25*(1), 7–22. https://doi.org/10.1023/A:1007580217973

Morales, E. E., & Trotman, F. K. (2004). *Promoting Academic Resilience in Multicultural America.*

Mornane, A. (2009). The adolescent students views of factors influencing their learning.

International Journal of Learning, *16*(5), 221–230. Retrieved from http://www.scopus.com/inward/record.url?eid=2-s2.0-77950199154&partnerID=40&md5=ffe2d07c058a4afca825eee81d0be4e7

Muratori, P., Bertacchi, I., Giuli, C., Nocentini, A., Ruglioni, L., & Lochman, J. E. (2016). Coping Power Adapted as Universal Prevention Program: Mid Term Effects on Children's Behavioral Difficulties and Academic Grades. *The Journal of Primary Prevention*, *37*(4), 389–401. https://doi.org/https://dx.doi.org/10.1007/s10935-016-0435-6

National Center for Educational Statistics. (n.d.). Science (2003). 15 years. Retrieved from https://nces.ed.gov/surveys/pisa/idepisa/report.aspx

Nota, L., Soresi, S., & Zimmerman, B. J. (2004). Self-regulation and academic achievement and resilience: A longitudinal study. *International Journal of Educational Research*, *41*(3), 198–215. https://doi.org/10.1016/j.ijer.2005.07.001

O'Neill, J., Tabish, H., Welch, V., Petticrew, M., Pottie, K., Clarke, M., … Tugwell, P. (2014a). Applying an equity lens to interventions: Using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. *Journal of Clinical Epidemiology*, *67*(1), 56–64. https://doi.org/10.1016/j.jclinepi.2013.08.005

O'Neill, J., Tabish, H., Welch, V., Petticrew, M., Pottie, K., Clarke, M., … Tugwell, P. (2014b). Applying an equity lens to interventions: Using PROGRESS ensures consideration of socially stratifying factors to illuminate inequities in health. *Journal of Clinical Epidemiology*, *67*(1), 56–64. https://doi.org/10.1016/j.jclinepi.2013.08.005

OECD. (2012). *Programme for international student assessment (PISA) results from PISA 2012*. *OECD*. https://doi.org/08/03/2017

Ogbu, J. (1991). Minority coping responses and school experience. Journal of. *Journal of Psychohistory*, (18), 433–456.

Ogbu, J. U., & Simons, H. D. (1998). Voluntary and Involuntary Minorities: A Cultural-Ecological Theory of School Performance with Some Implications for Education. *Anthropology Education Quarterly*, *29*(2), 155–188. https://doi.org/10.1525/aeq.1998.29.2.155

Ogilvie, D., Fayter, D., Petticrew, M., Sowden, A., Thomas, S., Whitehead, M., & Worthy, G. (2008). The harvest plot: A method for synthesising evidence about the differential effects of interventions. *BMC Medical Research Methodology*, *8*. https://doi.org/10.1186/1471-2288-8-8

Orfield, G. (2004). *Dropouts in America: Confronting the Graduation Rate Crisis*. Cambridge, MA, US: Harvard Education Press.

Orfield, G., & Lee, C. (2005). Why Segregation Matters: Poverty and Educational Inequality, 48. Retrieved from https://escholarship.org/content/qt4xr8z4wb/qt4xr8z4wb.pdf

Osborne, J. W., & Walker, C. (2006). Stereotype threat, identification with academics, and withdrawal from school: Why the most successful students of colour might be most likely to withdraw. *Educational Psychology*, *26*(4), 563–577. https://doi.org/10.1080/01443410500342518

Oxman, A. D., & Guyatt, G. H. (1992). A consumer's guide to subgroup analyses. *Ann Intern Med*, *116*(1), 78–84. https://doi.org/10.7326/0003-4819-116-1-78

Oxman, A., & Guyatt, G. (2002). When to Believe a Subgroup Analysis. *Users' Guides Interactive*, 251–262. Retrieved from http://www.usersguides.org.

Padrón, Y. N., Waxman, H. C., Brown, A. P., & Powers, R. A. (2000). *Improving classroom instruction and student learning for resilient and non-resilient English language learners*.

Pajares, F. (2001). Toward a positive psychology of academic motivation. *Journal of Educational Research*, *95*(1), 27–35. https://doi.org/10.1080/00220670109598780

Parker, S. W., Rubalcava, L., & Teruel, G. (2005). Schooling Inequality and Language Barriers. *Economic Development and Cultural Change*, *54*(1988), 71–94. https://doi.org/10.1086/431257

Peske, H. G., & Haycock, K. (2006). Teaching Inequality: How Poor and Minority Students Are Shortchanged on Teacher Quality: A Report and Recommendations by the Education Trust. *Education Trust*. Retrieved from http://eric.ed.gov/?id=ED494820

Petticrew, M., & Roberts, H. (2003). Evidence, hierarchies, and typologies: Horses for courses. *Journal of Epidemiology and Community Health*. https://doi.org/10.1136/jech.57.7.527

Petticrew, M., Tugwell, P., Kristjansson, E., Oliver, S., Ueffing, E., & Welch, V. (2011). Damned if you do, damned if you don't: Subgroup analysis and equity. *Journal of Epidemiology and Community Health*, *66*(1), 95–98. https://doi.org/10.1136/jech.2010.121095

Petticrew, M., Tugwell, P., Kristjansson, E., Oliver, S., Ueffing, E., & Welch, V. (2012). Damned if you do, damned if you don't: Subgroup analysis and equity. *Journal of Epidemiology and Community Health*, *66*(1), 95–98. https://doi.org/10.1136/jech.2010.121095

Polite, V. C., & Davis, J. E. (1999). *African American males in school and society: Practices & policies for effective education*. New York: Teachers College Press.

Powell, J. J. W. (2003). Constructing Disability and Social Inequality Early in the Life Course: the case of special education in Germany and the United States. *Disability Studies Quarterly*, *23*(2), 57–75.

RevMan. (2014). Review Manager (RevMan) [Computer program]. Copenhagen: The Cochrane Collaboration, The Nordic Cochrane Center.

Rigsby, L. C., Stull, J. C., & Nancy, M.-K. (1997). Determinants of student educational expectations and achievement: Race/ethnicity and gender differences. In *Social and emotional adjustment and family relations in ethnic minority families*. (pp. 201–224). Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1997-08310-011&site=ehost-live

Riley, J. R., & Masten, A. S. (2005). Resilience in context. *Resilience in Children, Families, and Communities: Linking Context to Practice and Policy*, 13–25. https://doi.org/10.1007/0-387-23824-7_2

Ross, S., Sanders, W., Wright, S. P., Stringfield, S., Wang, L. W., & Alberg, M. (2001). Two-

and Three-Year Achievement Results From the Memphis Restructuring Initiative. *School Effectiveness and School Improvement*, *12*(3), 323–346. https://doi.org/10.1076/sesi.12.3.323.3451

RStudio Team. (2015). RStudio: Integrated Development for R. Boston, MA: RStudio, Inc.

Rutter, M., Maughan, B., Mortimore, P., Ouston, J., & Smith, A. (1979). *15,000 hours: Secondary schools and their effects on children*. Shepton Mallet: Open Books.

Sailor, W. (2009). *Making RTI Work. Making RTI Work: How Smart Schools are Reforming Education through Schoolwide Response-to-Intervention*. https://doi.org/10.1002/9781118269480

Schünemann, H. J., Oxman, A. D., Vist, G. E., Higgins, J. P., Deeks, J. J., Glasziou, P. P., & Guyatt, G. H. (2008). Interpreting Results and Drawing Conclusions. In *Cochrane Handbook for Systematic Reviews of Interventions: Cochrane Book Series* (pp. 359–387). https://doi.org/10.1002/9780470712184.ch12

Schunk, D. H., & Zimmerman, B. J. (1998). *Self-regulated learning: From teaching to self-reflective practice. Psychological Science*.

Seligman, M. E. (1972). Learned helplessness. *Annual Review of Medicine*, 207–412. https://doi.org/10.1146/annurev.me.23.020172.002203

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). From: Shadish, W., Cook, T. & Campbell, D (2002) Experimental & Quasi-Experimental Designs for Generalized Causal Inference. Boston: Houghton Mifflin. *Experimental and QuasiExperimental Designs for Generalized Causal Inference*, *127*(2002), 64–102.

Shoshani, A., Steinmetz, S., & Kanat-Maymon, Y. (2016). Effects of the Maytiv positive psychology school program on early adolescents' well-being, engagement, and achievement. *Journal of School Psychology*, *57*, 73–92. https://doi.org/https://dx.doi.org/10.1016/j.jsp.2016.05.003

Snyder, F., Flay, B., Vuchinich, S., Acock, A., Washburn, I., Beets, M., & Li, K.-K. (2009). Impact of a social-emotional and character development program on school-level indicators of academic achievement, absenteeism, and disciplinary outcomes: A matched-pair, cluster randomized, controlled trial. *Journal of Research on Educational Effectiveness*, *3*(1), 26–55. https://doi.org/https://dx.doi.org/10.1080/19345740903353436

Steele, C. M. (1992). Race and the Schooling of Black Americans. *The Atlantic*. Retrieved from http://www.colorado.edu/ftep/sites/default/files/attached-files/ftep_memo_to_faculty_64.pdf%5Cnhttp://www.theatlantic.com/magazine/archive/1992/04/race-and-the-schooling-of-black-americans/306073

Steele, C. M. (1997). A threat in the air. How stereotypes shape intellectual identity and performance. *The American Psychologist*, *52*(6), 613–629. https://doi.org/10.1037/0003-066X.52.6.613

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

Sterne, J., Hernán, M., Reeves, B., Savović, J., Berkman, N., Viswanathan, M., … Altman,

D. (2016). The Risk Of Bias In Non-randomized Studies – of Interventions (ROBINS-I) assessment tooL.

Strand, J. A., & Peacock, T. D. (2002). Nurturing resilience and school success in American Indian and Alaska Native students. *EPIC Digest*. https://doi.org/Report No. RC 023771

Stromquist, N. P. (1990). Gender Inequality in Education: Accounting for women???s subordination. *British Journal of Sociology of Education*, *11*(2), 137–153. https://doi.org/10.1080/0142569900110202

Sun, X., Briel, M., Walter, S. D., & Guyatt, G. H. (2010). Is a subgroup effect believable? Updating criteria to evaluate the credibility of subgroup analyses. *BMJ (Online)*, *340*(7751), 850–854. https://doi.org/10.1136/bmj.c117

Swenson, L. M., Nordstrom, A., & Hiester, M. (2008). The Role of Peer Relationships in Adjustment to College. *Journal of College Student Development*, *49*(6), 551–567. https://doi.org/10.1353/csd.0.0038

Tak, Y. R., Kleinjan, M., Lichtwarck-Aschoff, A., & Engels, R. C. M. E. (2014). Secondary outcomes of a school-based universal resiliency training for adolescents: A cluster randomized controlled trial. *BMC Public Health*, *14*(1). https://doi.org/10.1186/1471-2458-14-1171

Tak, Y. R., Van Zundert, R. M., Kuijpers, R. C., Van Vlokhoven, B. S., Rensink, H. F., & Engels, R. C. (2012). A randomized controlled trial testing the effectiveness of a universal school-based depression prevention program "Op Volle Kracht" in the Netherlands. *BMC Public Health*, *12*, 21. https://doi.org/https://dx.doi.org/10.1186/1471-2458-12-21

Testa-Ryan, R. (2016). *Ecological Factors Promoting Academic Resilience of Latina Students*. Loyola University Chicago.

The Centre for Social Justice. (2014). *Closing the divide: Tackling educational inequality in England*. Retrieved from https://www.centreforsocialjustice.org.uk/library/closing-divide-tackling-educational-inequality-england

Tomasetto, C., Alparone, F. R., & Cadinu, M. (2011). Girls' math performance under stereotype threat: The moderating role of mothers' gender stereotypes. *Developmental Psychology*, *47*(4), 943–949. https://doi.org/10.1037/a0024047

Troyna, B. (1984). Fact or Artefact? The 'Educational Underachievement' of Black Pupils. *British Journal of Sociology of Education*, *5*(2), 153–166. https://doi.org/10.1080/0142569840050204

Troyna, B. (1991). Children, "race" and racism: The limitations of research and policy. *British Journal of Sociology of Education*, *39*(4), 425–436.

Troyna, B. (2012). *Racial Inequality in Education*.

Tugwell, P. (2006). Applying clinical epidemiological methods to health equity: the equity effectiveness loop. *BMJ*, *332*(7537), 358–361. https://doi.org/10.1136/bmj.332.7537.358

Tugwell, P., Petticrew, M., Kristjansson, E., Welch, V., Ueffing, E., Waters, E., … Kelly,

M. P. (2010a). Assessing equity in systematic reviews: realising the recommendations of the Commission on Social Determinants of Health. *BMJ*, *341*(sep13 1), c4739–c4739. https://doi.org/10.1136/bmj.c4739

Tugwell, P., Petticrew, M., Kristjansson, E., Welch, V., Ueffing, E., Waters, E., … Kelly, M. P. (2010b). Assessing equity in systematic reviews: Realising the recommendations of the Commission on Social Determinants of Health. *BMJ (Online)*, *341*(7778), 873–877. https://doi.org/10.1136/bmj.c4739

U.S. Department of Education. (2003). Indentifying and implementing educational practices supported by rigorous evidence: A user friendly guide. Retrieved from http://www2.ed.gov/rschstat/research/pubs/rigorousevid/rigorousevid.pdf

U.S. Department of Education. No Child Left Behind Act of 2001 (2011).

UNESCO. (n.d.). World Inequality Database on Education: Out-of-school adolescents. Retrieved November 12, 2017, from http://www.education-inequalities.org/indicators/edu_out_lowsec#?sort=mean&dimension=all&group=all&age_group=edu_out_lowsec&countries=all

UNESCO. (2013a). *Education for All Global Monitoring Report: Girls' Education - the facts*. Retrieved from https://en.unesco.org/gem-report/sites/gem-report/files/girls-factsheet-en.pdf

UNESCO. (2013b). *Inequalities in education; 2013*. Retrieved from http://unesdoc.unesco.org/images/0022/002204/220440E.pdf

UNESCO. (2014). *Shaping the Future We Want UN Decade of Education for Sustainable Development. 2005-2014 Final Report. United Nations Educational, Scientific and Cultural Organization*. https://doi.org/10.5363/tits.11.4_46

UNFPA. (2002). *State of the World Population: People, Poverty, and Possibilities*. Retrieved from http://www.unfpa.org/swp/2002/pdf/english/%0Aswp2002eng.pdf

United Nations. (2016). Goal 4 - Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all. *Sustainable Development Knowledge Platform*, 1–16. https://doi.org/sdg4

Valentine, C. A. (1968). *Culture and Poverty: Critique and Counter-Proposals*.

van Bergen, E., van Zuijen, T., Bishop, D., & de Jong, P. F. (2017). Why Are Home Literacy Environment and Children's Reading Skills Associated? What Parental Skills Reveal. *Reading Research Quarterly*, *52*(2), 147–160. https://doi.org/10.1002/rrq.160

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2015). Meta-analysis of multiple outcomes: a multilevel approach. *Behavior Research Methods*, *47*(4), 1274–1294.

Villanueva, I. (1996). Change in the educational life of Chicano families across three generations. *Education and Urban Society*, *29*, 13–34. https://doi.org/10.1177/0013124596029001003

Walberg, H. J. (1986). Home environment and school learning: Some quantitative models and research synthesis. *Child Rearing in the Home and School.*, 105–120. Retrieved from http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc2&NEWS=

N&AN=1986-98664-005

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1994). Educational resilience in inner cities. In *Educational Resilience in Inner-city America: Challenges and Prospects* (pp. 45–72).

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1997). Fostering educational resilience in Inner-City schools. *Children and Youth*, *7*, 119–140. Retrieved from http://www.temple.edu/LSS

Wang, W. (2005). Son preference and educational opportunities of children in China--- ``I wish you were a boy!". *Gender Issues*, *22*(2), 3–30. https://doi.org/10.1007/s12147-005-0012-4

Welch, V., Jull, J., Petkovic, J., Armstrong, R., Boyer, Y., Cuervo, L. G., … Tugwell, P. (2015). Protocol for the development of a CONSORT-equity guideline to improve reporting of health equity in randomized trials. *Implementation Science*, *10*(1), 1. https://doi.org/10.1186/s13012-015-0332-z

Welch, V., Petticrew, M., Petkovic, J., Moher, D., Waters, E., White, H., … White, H. (2016). Extending the PRISMA statement to equity-focused systematic reviews (PRISMA-E 2012): explanation and elaboration. *Journal of Development Effectiveness*, *8*(2), 287–324. https://doi.org/10.1080/19439342.2015.1113196

Welch, V., Petticrew, M., Tugwell, P., Moher, D., O'Neill, J., Waters, E., … White, H. (2012). PRISMA-Equity 2012 Extension: Reporting Guidelines for Systematic Reviews with a Focus on Health Equity. *PLoS Medicine*, *9*(10). https://doi.org/10.1371/journal.pmed.1001333

Welch, V., Petticrew, M., Ueffing, E., Jandu, M. B., Brand, K., Dhaliwal, B., … Tugwell, P. (2012). Does consideration and assessment of effects on health equity affect the conclusions of systematic reviews? a methodology study. *PLoS ONE*, *7*(3), 1–6. https://doi.org/10.1371/journal.pone.0031360

Welch, V., Tugwell, P., & Morris, E. B. (2008). The equity-effectiveness loop as a tool for evaluating population health interventions. *Revista de Salud Publica (Bogota, Colombia)*, *10 Suppl*(1), 83–96. https://doi.org/10.1590/S0124-00642008000600008

Welch, V., Tugwell, P., Petticrew, M., J, de M., Ueffing, E., Kristjansson, B., … Smylie, J. (2010). How effects on health equity are assessed in systematic reviews of interventions. *Cochrane Database of Systematic Reviews*, N.PAG-N.PAG. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=104970636&site=ehost-live

Werner, E. E., & Smith, R. S. (1992). Overcoming the odds: High risk children from birth to adulthood. *Overcoming the Odds: High Risk Children from Birth to Adulthood.*

Westen, D., & Rosenthal, R. (2005). Improving construct validity: Cronbach, Meehl, and Neurath's ship. *Psychological Assessment*, *17*(4), 409–412. https://doi.org/10.1037/1040-3590.17.4.409

Whitehead, M. (1992). The Concepts and Principles of Equity and Health. *International Journal of Health Services*, *22*(3), 429–445. https://doi.org/10.2190/986L-LHQ6-2VTE-YRRN

WHO. (2007). A Conceptual Framework for Action on the Social Determinants of Health.

COMMISION ON SOCIAL DETERMINANTS OF HEALTH (DRAFT).

World Bank. (2016). World Bank: Country and Lending Groups | Data. *World Bank Data Website*, 1–8. Retrieved from http://databank.worldbank.org/data/download/site-content/CLASS.xls

Zimmerman, B., & Schunk, D. (2001). *Self-Regulated Learning and Academic Achievement: Theoretical Perspectives*. Retrieved from https://books.google.com/books?id=og4hVOcjcqMC&pgis=1