Analyses of Social Emotional Learning Outcomes in the Age of Diversifying Populations and Locally Developed Programs

A DISSERTATION

SUBMITTED TO THE FACULTY OF

THE UNIVERSITY OF MINNESOTA

BY

Anna Li, M.A.

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Advisor: Dr. Faith G. Miller

May 2025

Acknowledgments

Although my name is listed as the author of this dissertation, it should be known that a dissertation, like most things in life, is never accomplished alone. My work, first and foremost, would not be possible without my advisor Dr. Faith Miller who has guided me through every step of my graduate school journey. Dr. Miller, I cannot thank you enough for taking a chance on accepting into this Ph.D. program a 22-year-old, first-generation student, and not-even-out-of-undergrad individual who felt she could barely string together two nervous thoughts on interview day. Your belief in me, academic wisdom, and thoughtful encouragement throughout these past four years have truly strengthened and shaped me into the scholar I am today. To my committee chair, Dr. Lisa Aguilar, and my committee members Dr. Annie Hansen-Burke and Dr. Kim Gibbons, thank you all for your valuable feedback and constructive thoughts that have helped build and clarify this body of work. The perspectives you all brought to this project allowed me to deeply consider aspects of this research area that I had not previously considered.

To my 爸爸 and 妈妈, thank you for trusting all of your children to honestly and passionately pursue the paths of education that interested us. Your belief in our decisions has encouraged us to be the strong individuals we are today. To my siblings Anne, Angie, and Andy, thank you for supporting me in your own, unique, ways. I would also like to thank all of my friends who have been so supportive throughout this process. To Abigail, I'm thankful for our numerous data collection journeys that allowed for mini adventures throughout the year. Greta, thank you for last minute working with me; Gchatting with you made the most painful portion of editing and finalizing so much more enjoyable. To Shay, Pete, and Elizabeth, I credit making it through graduate school because of the strength and humor you all provided to our doc cohort. I also give huge thanks to my friend Emma for

working with me throughout the summer for accountability writing. I simply could not have completed this dissertation working without your company.

Next, a very special shout out to my cat Tofu who has reminded (read: demanded) me to take breaks from my computer screen so I can get up and play with her. You remind me each day that there is more to life than work. Lastly, I would like to thank my partner Jared who has been by my side from applying to graduate school, moving to a new state, and throughout these past four years. You have been my rock, voice of reason, and the biggest supporter I could have ever dreamed of having by my side. Your belief in and care for me has gotten me through my darkest days, and I will forever be grateful you willingly choose to spend every day with me. I would be drowning in the dishes I don't do without you.

Dedication

I dedicate this work to the adults who, as young students, so bravely and unnecessarily navigated education systems that failed to support their academic, social, emotional, and behavioral development simply because they did not belong to a majority culture. This work is for all the adults with the inner child who was made to feel "lesser than" or "foreign" because they *dared* to be born into a system that could not box their identities into oppressive categories. To all the future generations of young students going through our broken and unjust school systems, I will dedicate my work and continue the fight to ensure school feels like a safe and welcoming place for you to learn and flourish like you so deserve.

Abstract

The complexity of the world in which students are developing seems to increase with each unprecedented event that occurs in the United States and across the globe. Implementing social emotional learning (SEL) programs is one way in which PreK-12 schools can support students in learning the necessary skills to not only navigate personal and interpersonal relationships, but also work toward promoting equity to make meaningful changes before, during, and after these unprecedented events (CASEL-a). However, even universal SEL programs may not necessarily serve all students. Thus, this two-study dissertation examined student outcomes in regard to SEL programs under different circumstances. The first study was a systematic review of the empirical literature where authors synthesized the evidence-base of cultural adaptations to SEL programs. The second study was a preliminary analysis of student outcomes in regard to a locally developed SEL program called Be Good People (BGP). Results of both studies were mixed. Findings from the first study demonstrated there were remarkably few studies that examined the impact of cultural adaptations on student outcomes in SEL programs. There were mixed results in terms of the effectiveness of the cultural adaptations on various student outcomes. However, there were high levels of acceptability and feasibility for culturally adapting the SEL programs. Similarly, the second study found high practicality levels for implementing BGP. Teacher ratings also indicated some increased SEL skills and academic performance Time 1 to Time 2 in the study. However, the program demonstrated a lack of cultural relevance for diverse student backgrounds. Given these mixed findings, the intersection of culturally relevant pedagogy and SEL programs is an area necessitating future research. Other implications for research and practice are discussed.

Table of Contents

List of Figures

List of Tables

List of Appendices

Positionality

When I read research, I work to remember that everyone has their own intersecting identities that guide their worldview. As an individual now engaging in research, I am no different—I bring my experiences as a first-generation student, cis-gendered, heterosexual, able-bodied, Chinese American woman to every step of research I do. My interest in social emotional learning (SEL) stems from the context in which I was raised, the positions I have held, and the political society in which we exist. Growing up in an immigrant household in Milwaukee, WI suburbia meant I was often in the middle of two cultures. Existing in the middle of two worlds, I noticed there were people uncomfortable around unfamiliar things, situations, and people. I saw the best and worst of individuals when it came to acceptance of others based on the nature of the environment in which they grew. This experience growing up undoubtedly influenced why I chose to conduct my research in the SEL realm.

Further, the dedication I have written for this study includes myself. This work has been, in part, a personal journey in recognizing the culturally relevant pedagogy I failed to receive in the classrooms where I learned. However, I firmly believe that the emotion I bring to my search for how SEL programs can be used as a tool for equity-centered practices in classrooms is a step toward recognizing that strong science is not just cold numbers, rather strong science leverages numbers defined with humanity and intentionality—both of which require strong emotion.

In the first study, Chapter 2 of this dissertation, I take an equity-centered lens in which I critically examine the extent to which SEL programs are suitable for *all* students through adaptations that are made for culturally diverse students. In the second study, Chapter 3 of this dissertation, I further grounded my study in strong, school-research partnerships. This meant demonstrating interest and including research questions centering what our district contact, the SEL coordinator, desired from this study. Moreover,

strong school-research partnerships, in my view, required respecting the school in which I, as a researcher, was privileged enough to be welcomed. For me, respect was shaped by values with which I was raised. For example, I simply could not burden teachers with extensive measures and an expectation of high fidelity. Entering a school-research partnership for me meant balancing strong scientific practices with the recognition that our project would be far from perfection because we were working within the stressed education system.

Finally, I recognize that the purpose for this research was fueled by a desire to see **my** version of a more just world, which has influenced the questions I have asked, methods I've leveraged, discussions I have written, and implications for future research and practice I have posed. I believe the purpose of research is to act as a vessel to this more just world, and my intersecting identities and experiences have shaped my view of what I believe to be important in the world.

Chapter 1

Introduction

With the majority of students' waking hours spent in class, school is a crucial place where they develop peer relations and learn essential social and emotional skills (Hall & Nielsen, 2020). Students can learn these skills in a multitude of ways, but there is a growing emphasis on directly teaching these skills in schools through the implementation of social and emotional learning (SEL) programs (Tyton Partners, 2021). Indeed, SEL has shown benefits for student improvements in academics and social, emotional, and behavioral functioning (Cipriano et al. (2023); Durlak et al., 2011). Despite these benefits, SEL programs are not immune to flaws. Like many other educational initiatives in the United States, it is imperative that SEL programs that are broadly disseminated across whole classrooms equitably serve all students. Educators have expressed concern that current SEL practices can further harm Black, Indigenous, and people of color (BIPOC) youth, where SEL could unintentionally strip youth of their cultural identities (Lin et al., 2023). Thus, engaging in SEL with a critical lens is critical when evaluating the landscape of SEL.

The intersections of equity and SEL have many facets that can be analyzed including, but not limited to, program construction, settings in which programs are implemented, and student populations programs are designed to serve. An expansive meta-analysis from Cipriano and Colleagues (2023) examined 424 studies from over 50 countries, and yet, still discussed how the SEL field could benefit from more information regarding the intersection of present SEL knowledge and culturally based frameworks (e.g. culturally relevant pedagogy). With the expanding diversity of student identities in United States' schools, it's crucial to design and deliver effective SEL programs to students with varying backgrounds, languages, and identities. Given the lack of knowledge in culturally based frameworks and SEL, the purpose of this two-study dissertation is to elucidate the

relationship between SEL and student outcomes through (1) systematically reviewing the landscape of cultural adaptations to SEL programs and (2) exploring an up-and-coming, widely disseminated, locally developed SEL program called Be Good People, for its relation to supporting positive student outcomes.

Background and Rationale

The need for social emotional support is becoming increasingly necessary for students in schools. In recent history, effective social and emotional skills have been necessary for students to navigate a shift from in-person school to online school and back to in-person school due to the COVID-19 pandemic (Darling-Hammond et al., 2020). The disruption in typical routine the COVID-19 pandemic has had on students' social, emotional, and behavioral well-being has been markedly high. School systems have experienced increases in mental health issues and problem behaviors (NCES, 2022), and lower academic achievement (Fisher et al., 2022). Thus, the past few years have made clear the necessity for all students to gain social and emotional skills.

Universal SEL programs are intended to serve *all* students. Skills learned from universal SEL programs could include problem solving, effective communication, building healthy relationships, inter- and intra-personal emotion identification and management, and working with people from different backgrounds (CASEL, n.d.-a). Many of these skills are used in school settings where students develop academically, socially, emotionally, and behaviorally. Although, universal SEL programs have been shown to be beneficial in some regards to student outcomes (Cipriano et al., 2023), the positionality of: (a) university faculty who are frequently developing and researching SEL programs and (b) teachers who are implementing these programs should be taken into consideration. When discussing "universality," the world views these individuals bring can vary by factors such as race. Notably, 73% of university faculty are white (NCES, 2023-b) and 80% of K-12 educators are

also white (NCES, 2023-a). In contrast, the student population in 2021 K-12 public schools was only 45% white, with that percentage expected to fall to 42% in 2031 (NCES, 2023-c).

Some universal SEL programs have found their curricula suitably serve students from a variety of backgrounds (e.g., Positive Action program; Duncan et al., 2016, LaBelle, 2019). However, through a historical lens, many universal SEL programs have been developed based on white, middle-class norms and have been implemented to serve white, middle-class students (Clark et al., 2022; Hoffman, 2009; Mahfouz & Anthony-Stevens, 2020). These demographics are no longer the norm in many classrooms across the United States (NCES, 2023). Culturally relevant pedagogy as a theoretical basis for any curriculum must provide a way for students to both succeed academically while maintaining their cultural identities and learning about other cultural identities (Ladson-Billings, 1995). SEL programs could be an effective avenue for students to learn crucial skills to navigate an ever-growing and complex world, however, there is concern for whether SEL programs are culturally relevant (Mahfouz & Anothony-Stevens, 2020). An educational initiative situated in a culturally relevant pedagogy framework promotes students' academic achievement by upholding their cultural identities and encouraging students to identify, analyze, and problem solve societal inequities (California Department of Education, 2023). It is therefore imperative for researchers to explore the relationships between universal SEL programs and student outcomes for students from multicultural backgrounds and investigate the effectiveness of widely implemented SEL programs. Thus, this dissertation undertakes two studies to add to the SEL practices knowledge base.

Study 1

Although calls have been made for greater efforts to provide culturally relevant education to better serve all students (Bennouna et al., 2021), application of these efforts to the domain of SEL remain somewhat limited. One way that SEL researchers have attempted

to provide more culturally relevant SEL programs has been through making cultural adaptations to existing SEL programs. There has not yet been a systematic review of the literature examining cultural adaptations to SEL programs in PreK-12 schools. Consequently, Study 1 undertook a systematic review of the empirical literature to better understand the adaptations being made, the purposes of the adaptations, and the association of the adaptations with student outcomes. The research questions for Study 1 were as follows:

1. What types of cultural adaptations have been made to SEL programs in the empirical literature?
2. What are the stated purposes of the cultural adaptations made to SEL programs in the empirical literature?
3. To what extent are these cultural adaptations of SEL programs associated with student outcomes?

Study 2

An educator developed, open-access, SEL program titled Be Good People (BGP) is freely available to use. Districts across the country and on all continents but Antarctica have registered to use the program. However, no research has been conducted on the program, which was created in 2020. Thus, Study 2 was a feasibility study of the BGP program, designed to gain preliminary evidence of outcomes associated with the program. The research questions for Study 2 were as follows:

1. To what extent do teachers implement BGP modules with fidelity (ie., dosage and adherence)?
2. Are improved student social emotional skills observed over six months of implementation?
3. Are improved academic and behavioral student outcomes observed over six months

of implementation (i.e., grades and social emotional behavioral screening data)?

4. What are teacher perceptions of the practicality of BGP?
    a. Practicality for this study was defined as the capability to implement the intervention (Bowen et al., 2009).
5. What are teacher perceptions of the acceptability of BGP?
    a. Acceptability for this study was defined as the suitability and satisfaction of the intervention (Bowen et al., 2009).
6. What are perceptions of the cultural relevance of BGP?
7. What changes do teachers believe should be made to BGP (if any)?
8. What changes do teachers believe should be made to BGP (if any)?

Summary

It is broadly understood that SEL programs are helpful in boosting student academic and behavior outcomes. However, more information is needed to better understand the following: (1) how existing SEL programs are being made more culturally relevant through adaptations and (2) how newly developed SEL programs are serving teachers and students, both in terms of student outcomes and cultural relevance. The two studies in this dissertation lay out a contemporary understanding of student outcomes associated with SEL programs, with a particular focus on cultural relevance.

Key definitions

- *Social emotional learning (SEL):* Conceptualized as the process in which children, adolescents, and adults learn knowledge and skills to achieve healthy identity development, emotion management, and goal achievement (CASEL, n.d.-a),
- *Culturally Relevant Pedagogy*: This theoretical model aims to support student academic learning, learning of their own and others' cultures, and developing a sense of critical consciousness (Ladson-Billings, 1995). Related terms include

"culturally responsive teaching" and "culturally sustaining pedagogy." Both terms stem from culturally relevant pedagogy. Culturally responsive teaching refers to leveraging cultural background, knowledge, and experiences of students to make learning more effective (Gay, 2018). Culturally sustaining pedagogy understands that students operate in a pluralistic cultural society and seeks to uphold (i.e., sustain) their own cultural identities while also having access to learning the dominant culture (Paris, 2012). This dissertation will focus on culturally relevant pedagogy as it is the foundational work to the related terms.

- *Critical Consciousness:* learning to perceive social, political, and economic contradictions, and to take action against the oppressive elements of reality (Freire, 1970, p. 35).
- *Culture:* "A socially transmitted or socially constructed constellation consisting of such things as practices, competencies, ideas, schemas, symbols, values, norms, institutions, goals, constitutive rules, artifacts, and modifications in the physical environment" (Fiske, 2002, p. 85).
- *Cultural adaptations:* the "systematic modification of an evidence-based treatment (EBT) or intervention protocol to consider language, culture, and context in such a way that is compatible with the client's cultural patterns, meanings, and values" (Bernal et al., 2009, p. 362).
- *Be Good People (BGP)*: A widely disseminated, open access, universal SEL curriculum at its Tier 1 level that was developed in April 2020 by educators in the St. Croix River Education District in Minnesota.

Chapter 2

Study 1: Cultural Adaptations to Social Emotional Learning Programs: A Systematic Review

Abstract

Social emotional learning (SEL) programs offer critical opportunities to promote the development of students' social awareness and emotional well-being. Yet, the landscape of how educators and researchers are making these programs effectively serve all students is unknown. The purpose of this study was to investigate cultural adaptations to SEL programs. This study systematically reviewed the types and stated purposes of cultural adaptations to SEL programs and the extent to which they are associated with student outcomes in the empirical literature. Included studies were (a) peer-reviewed empirical studies or dissertations; (b) conducted in Prekindergarten (PreK) through 12th grade school settings in the United States; (c) available in English; and (d) included a cultural adaptation to an SEL program. Five electronic databases were searched in January 2023. Included articles were coded to extract information regarding the types, purposes, and associations with student outcomes and cultural adaptations to SEL programs. A total of 11 studies, including 5,173 students, met inclusion criteria. Results demonstrated most studies used surface and deep level adaptations for the purpose of adapting SEL programs to a specific racial/ethnic group and/or a geographical region. In general, there were mixed results in terms of effectiveness when examining effect sizes and other statistical analyses of the adaptations on student outcomes. However, most studies found high social validity in relation to the adaptations. Implications and directions for future research are discussed.

Keywords: social emotional learning, adaptations, equity

Cultural Adaptations to Social Emotional Learning Programs: A Systematic Review

Schools are one of the first places many students learn how to socialize, empathize, and problem solve with their peers. Skills such as taking turns, listening, and managing big feelings are not only necessary in schools, but also in adulthood. These are examples of skills that fall under the umbrella of social emotional learning (SEL). How students learn these skills can influence how they behave towards and impact others. Importantly, students become adults who make decisions on how society functions, making learning these skills key in child development. One growing method schools have taken to teach SEL is through the use of social emotional learning programs. In a national survey of over 1000 teachers' and principals' SEL implementation, 76% of principals and 53% of teachers reported using SEL programs during the 2021-2022 school year—a 25% growth from 2018 (Schwartz et al., 2022). With expanding use and the increasing diversity in the United States', an examination of how the educational programs aligns with the needs of all students is critical.

Culturally relevant pedagogy and the field of social emotional learning (SEL) are currently growing areas of research in schools (Jones & Doolittle, 2017); however, there is relatively little known about how the two research areas intersect. Yet, the need is clear: in 1995, white students made up 65% of the population in public elementary and secondary schools, compared to the expected 45% in 2029 (NCES; National Center for Education Statistics, 2022). Despite this growth in diversity, SEL programs in K-12 schools have centered on westernized, middle-class, white norms (Hoffman, 2009; Kaler-Jones, 2020). This form of SEL is no longer suitable for the majority of students (Jagers et al., 2019). As such, advocates for SEL—known as the process in which people learn to show empathy, establish relationships, and make thoughtful decisions (Collaborative for Academic, Social,

and Emotional Learning; CASEL, n.d.-a)—have noted a need to move toward more inclusive and equitable programs (Hoffman, 2009; Jagers et al. 2021).

The degree to which SEL programs have been adapted to encompass culturally relevant pedagogy is not currently known. Although there is some evidence that researchers have made adaptations to SEL programs (e.g., Castro-Olivio & Merrell, 2012; Hopson, 2006), researchers have yet to systematically review the extant literature to better understand the types of adaptations being made to these SEL programs and associated outcomes. As such, the aim of this systematic review was to detail the types of cultural adaptations performed in empirical literature within Pre-kindergarten (PreK) through grade 12 (PreK-12) schools to SEL programs and to better understand the extent to which they support culturally diverse students.

Social Emotional Learning

The term SEL was coined in 1994 to address the "missing piece" in education to fulfill the ideal of a whole child approach (CASEL, n.d.-b), and in 2004, Illinois became the first state to adopt K-12 SEL standards (Dusenbury and Weissberg, 2017). Conceptualized as the process in which children, adolescents, and adults learn knowledge and skills to achieve healthy identity development, emotion management, and goal achievement (CASEL, n.d.-a), the adoption of SEL programs across the U.S. has now become more widespread. A 2018 systematic review examining states with freestanding and integrated SEL standards found that all 50 U.S. states had freestanding preschool standards, 49 states had K-12 SEL competencies integrated into existing health and physical education standards, and 11 states had freestanding K-12 SEL education standards (Eklund et al., 2018).

Teaching SEL in schools through a whole-child approach has demonstrated benefits (Elbertson et al., 2010). SEL programs have promoted growth in SEL skills, attitudes, behaviors, and academic performance across thousands of children (Durlak et al., 2011).

Programs that teach SEL leverage direct instruction and practice to develop relevant skills (Durlak et al., 2011). A report summarizing three large-scale reviews found that SEL programs promoted an average of 11 to 17 percentile point gain on standardized achievement tests, decreased conduct problems, emotional distress, and increased social-emotional skills, such as having a more positive attitude toward oneself (Payton et al., 2008). The benefits of SEL programs can be likened to those of prevention-focused public health programs, where the investment in programming can improve the well-being of the general public (Greenberg et al., 2017); however, similar to current challenges in public health (Devakumar et al., 2020), racism and systems of oppression undermine these efforts.

Many SEL programs in the United States centralize white, middle-class values which ignore and devalue many students' cultures, beliefs, and norms (Clark et al., 2022; Hoffman, 2009; Mahfouz & Anthony-Stevens, 2020). In general, this pattern is consistent with the broader educational environment. Textbooks in the U.S. tend to center themes of individuality (e.g. self-direction, achievement) over collectivism (e.g. conformity, group harmony) (Imada, 2010). Teaching all students social-emotional norms that center on typical white, western values, (e.g. open self-expression) has several negative impacts: minoritized students can experience stress, along with student strengths being underestimated and undervalued by school staff (Tobin, 1995; Allen et al., 2013).

SEL programs tend to focus on the experiences of students from non-minoritized groups and what practical skills they can learn from a color-evasive perspective (Gregory & Fergus, 2017), meaning that student racial and ethnic identities are ignored to support uniformity among students (Ward Schofield, 2019). Non-inclusive (i.e., color-evasive), white-centered norms, expectations, and practices that are not critically examined, can contribute to more harm than support for minoritized student populations. For example, when disaggregating Asian American and Pacific Islander groups, Pacific Islanders were

twice as likely as their white peers to be at risk for school discipline (Nguyen et al., 2019). Further, when looking at behavioral consequences, Black students at the secondary level in the 2011-2012 academic year were, nationally, had a suspension rate that was 16.5% higher than those of white students (Losen et al., 2015). The cultural mismatch hypothesis posits that Black students learning from curricula, teachers, and expectations coded with white, middle-class norms can contribute to the discipline gap (Monroe, 2005).

Moreover, historically, much of SEL discourse has been held by white researchers with much programming targeting Black and Brown youth and their behaviors (Kaler-Jones, 2020; Starr, 2019). An example of this phenomenon is "grit," which was conceptualized as perseverance and consistency toward a goal (Credé, 2018, Datu & Restubog, 2020; Hatchimonji, 2016). Looking to Black and Brown youth to have "grit" in the face of adversity, or worse—questioning if they do—is meaningless and harmful, particularly when scholars do not recognize history and implement true systemic changes that remove barriers for Black and Brown kids' success (Love, 2019). Removing barriers and preventing systemic harm includes ensuring that the programs students learn from is not the current unrelatable, irrelevant, and inherently racist curriculum that is often seen in the classroom (Davis, 2018; Klees, 2017; Labaree, 2009)

Culturally Relevant Pedagogy (CRP)

A culturally relevant education is of utmost importance to serve all students effectively & equitably (Brown-Jeffy & Cooper, 2011). Proposed in 1995, Dr. Gloria Ladson-Billings described the theoretical underpinnings of Culturally Relevant Pedagogy theory: conceptions of self and others, social relations, and conceptions of knowledge. These underpinnings arose from a study of eight teachers who were exceptional at supporting Black students (Ladson-Billings, 1995). Culturally relevant pedagogy (CRP) focuses on increasing student learning, cultural humility, and critical consciousness to support student

achievement and challenge societal inequities while centering students' identities (California Department of Education, 2023; Ladson-Billings, 1995). Brown-Jeffy and Cooper (2011) developed a conceptual framework of CRP based on principles from and foundational work regarding Critical Race Theory (Ladson-Billings & Tate, 1995) and multicultural (Gay, 1994; Nieto, 1999), culturally responsive (Gay, 2018), and effective (Ladson-Billings, 1994) learning and teaching. These principles encompassed identity and achievement, equity and excellence, developmental appropriateness, teaching the whole child, and student teacher relationships (Brown-Jeffy & Cooper, 2011).

Ladson-Billings (1995) described exemplary teachers of Black students as supporting their conceptions of self and others through believing all students were capable of success. These teachers also saw their pedagogy as always changing and improving and saw themselves as members of a larger community where teaching was a way to give back (Ladson-Billings, 1995). Dr. Ladson-Billings also noted in her seminal proposal of the Culturally Relevant Pedagogy theory that teachers who upheld social relations kept flexible student-teacher relations, were connected with their students, developed a community for learners, and pushed students to collaborate with each other. Lastly, in her investigation, the ways in which teachers viewed knowledge arose as a large part of the theory of Culturally Relevant Pedagogy. Exceptional teachers in the study understood that knowledge is shared, should be questioned, and that they must demonstrate a passion for learning. This critical approach and passion for knowledge and learning was then scaffolded to facilitate deeper learning among students. As such, culturally relevant pedagogy frames the work in the present study to ensure that SEL programs are relevant for students.

Cultural Adaptations

Culture is an incredibly complex and nuanced construct. Fiske (2002) examines the interconnecting aspects of culture across the world and defines it as "a socially transmitted

or socially constructed constellation consisting of such things as practices, competencies, ideas, schemas, symbols, values, norms, institutions, goals, constitutive rules, artifacts, and modifications in the physical environment" (p. 85). Bernal and colleagues (2009) defined cultural adaptations as the "systematic modification of an evidence-based treatment (EBT) or intervention protocol to consider language, culture, and context in such a way that is compatible with the client's cultural patterns, meanings, and values" (p. 362). They further describe cultural adaptations as the middle ground between creating new treatments for specific cultural groups (i.e., cultural compatibility of treatments) and testing current treatments across all cultural groups before making adaptations (i.e., universalistic hypothesis).

Increasing the cultural relevance of interventions has demonstrated positive effects in several research areas such as physical (Nierkens et al., 2013), behavioral (Barrera et al., 2013), mental health (Chowdhary et al., 2014), and education (Luevano & Collins, 2020). Cultural adaptations aid in making interventions more relevant for a more diverse student population. An example of this kind of cultural adaptation comes from a study by Slaten and colleagues (2015), which was conducted in an urban school where pedagogy at the school centered the student's unique locational (i.e., living in an urban setting) and racial/ethnic needs (e.g., a focus on critical consciousness and providing discussions of what it means to be Black) (Slaten et al., 2015).

Adaptations to programs can include, but are not limited to, changes to the language in which an intervention is delivered, the activities in a program, or the implementer. Cultural adaptations to programs may foster engagement among previously disengaged participants; target specific risk or resilience factors relevant to specific cultural groups; or replace ineffective interventions delivered to racially and ethnically diverse groups (Castro et al., 2010). A 2006 meta-analysis of culturally adapted mental health interventions found

that when adaptations were made toward the culture of the clients, they were four times more effective than when they were not. Additionally, when interventions were conducted in one's native language, they were twice as effective as just being conducted in English (Griner & Smith, 2006).

In accordance with the cultural sensitivity model (Resnicow et al., 1999), one method of categorizing cultural adaptations includes grouping them into surface and deep adaptation types. Surface structure features of cultural adaptations reflect the intervention presentation and deep structure features of cultural adaptations reflect the content of the intervention. Making cultural adaptations to interventions has demonstrated promise for delivering desirable outcomes. Given the importance and growth of SEL programming, as well as the highly interpersonal nature of it, consideration for the power of cultural adaptations to this area is important to examine. With the growth of SEL programs and diversity in the U.S., an examination of cultural adaptations to specifically SEL programs is essential.

Current Gaps in the Literature

Previous work has examined the impact of cultural adaptations for school-based social, emotional, and behavioral (SEB) interventions for students of color. One systematic review by Brown and colleagues (2018) found that cultural adaptations to SEB interventions (such as engaging outside partners) were used in studies with the intent of reducing exclusionary discipline practices on students of color. However, this study was broad in its scope (i.e. included all SEB interventions) yet narrow in the terms they used in their search: "cultural adaptation*," AND "social*," OR "emotional*," OR "behavior*" AND "intervention" OR "program" OR "practice" (e.g. "cultural adaptation" is a very specific term that could have other synonyms). There were also only three databases searched (i.e., ERIC, ProQuest Dissertations and Abstracts, and PsychINFO). Through this limited search, the

authors found that included studies used various adaptations related to adapting the content (e.g., language matching an intervention), procedures (e.g., engaging the community to provide input on the adaptations selected), and the program delivery (e.g., matching the cultural background of the interventionist with the recipients of the intervention). The researchers determined that these adaptation types are only the starting point of adapting SEB interventions. They noted the importance of determining specific school and student needs when making adaptations.

Moreover, calls have been made for adaptations to SEL using the interdisciplinary framework presented by Ladson-Billings (Mahfouz & Anthony-Stevens, 2020). Indeed, few PreK-5 SEL programs have a strong focus on equity-centered practices (i.e. lens of culturally relevant pedagogies, social justice-orientation, transformative SEL, and trauma-sensitivity) (Ramierez et al., 2021). Despite this gap in research, there are researchers who have examined how incorporating culturally relevant pedagogy and adaptations in their programs influence the outcomes on minoritized students.

One systematic review by McCallops and colleagues (2019) examined 51 studies that conducted SEL interventions in urban settings and found only five studies that provided evidence of relevant practices. Although this study shed light on the limited number of programs using culturally relevant pedagogy there was a narrow conceptualization of relevant settings excluding places such as suburban or rural settings. The team also leveraged a broad definition of culturally relevant pedagogy that included: empathy, self-awareness, awareness of others, self-regulation, and motivation. This broad definition does not aid in the specification of the types of cultural adaptations being made. Thus, a need for a broader lens on cultural adaptations is necessary.

Present Study

This study undertakes a systematic review of the empirical literature to better understand the evidence base surrounding types of cultural adaptations that have been made to SEL programs and how these adaptations are associated with student outcomes. Specific research questions were posed as follows:

1. What types of cultural adaptations have been made to SEL programs in the empirical literature?
2. What are the stated purposes of the cultural adaptations made to SEL programs in the empirical literature?
3. To what extent are these cultural adaptations of SEL programs associated with student outcomes?

Given the exploratory nature of these research questions, no hypotheses were made.

Method

Search Strategy

The authors used several methods to ensure a comprehensive review of the literature surrounding cultural adaptations to social emotional learning programs. In consultation with a university librarian, databases were selected if they were deemed relevant to psychology, education, and educational sciences. Databases were also selected if other key social emotional reviews or meta-analyses (e.g. Durlak et al., 2011) used similar databases in their search. Five electronic databases were searched in January 2023: Education Resources Information Center (ERIC) and Education Source (both via EbscoHost), PsychInfo and Medline (both via Ovid). The final database - Dissertations and Theses Global (via ProQuest) - was restricted to only dissertations as the dissertation process is similar to the peer review process, and presumably these studies should be of sufficient quality. No restrictions were placed on year within the databases.

Second, in consultation with the same university librarian, a list of key terms were first created. Search terms broadly covered the problem, intervention, and context of established research questions. Using the CASEL list of approved social emotional learning programs from October 2022 (CASEL, 2023), a list of programs and common acronyms for the programs was also created. The same search terms and strategies were used in each database. The searched terms included:

1. [Subject field code] Social Emotional Learning **OR**
2. [Abstract and title code] "social emotional learning" OR "social-emotional learning" OR "SEL" OR "social emotional competence" OR “social-emotional competence” **OR**
3. [Abstract and title code] “.b” OR “4Rs” OR “Adventures Abroad the S.S. Grin” OR “Al’s Pals” OR “BASE Education” OR “Building Assets, Reducing Risks” OR “BARR” OR “Captain McFinn” OR “Caring School Community” OR “Changemakers” OR “CHARACTERplus Way” OR “Circle of Education” OR “City Year’s Whole Child Whole School Model” OR “Competent Kids, Caring Communities” OR “Connect Science” OR “Conscious Discipline” OR “Consistency Management and Cooperative Discipline” OR “EduGuide” OR “EL Education” OR “Emotional ABCs Premium Schools” OR “Facing History and Ourselves” OR “Get Real” OR “Getting Along Together” OR “GiveThx” OR “Growing Early Mindsets” OR “GEM” OR “Hallway Heroes” OR “Harmony SEL” OR “Heartfulness Program for Schools” OR “HighScope Educational Approach for Preschool” OR “I Can Problem Solve” OR “Incredible Years” OR “Inner Explorer” OR “INSIGHTS Intervention” OR “Keepin it REAL” OR “Kindness in the Classroom” OR “Leader in Me” OR “Learning to Breathe” OR “Lion’s Quest” OR “Manners of the Heart” OR “Merrell’s Strong Kids” OR “Michigan Model for Health” OR “Mindful Practices and Class Catalyst” OR “MindUP” OR “Mosaic” OR “My Teaching Partner” OR “Open Circle” OR “Passport Skills for Life” OR “PBLWorks” OR

"Peace Learning Circles" OR "PeaceWorks" OR "Peer Group Connection" OR "Positive Action" OR "Promoting Alternative THinking Strategies" OR "PATHS" OR "Pure Power" OR "Raising Healthy Children" OR "Reading Apprenticeship" OR "Reading with Relevance" OR "Ready to Learn" OR "Resolving Conflict Creatively Program" OR "RCCP" OR "Responding in Peaceful and Positive Ways" OR "Responsive Classroom" OR "Ripple Effects for Teens" OR "RULER Approach" OR "School-Connect" OR "Second Step" OR "Sky Schools" OR "Smart Brain Wise Heart" OR "Social Decision-making" OR "Problem Solving Program" OR "Social Skills Improvement System Classroom Intervention Program" OR "SSIS" OR "SPARK" OR "Student Success Skills" OR "Teen Connection Project" OR "Teen Outreach Program" OR "The Character Effect" OR "The Fourth R" OR "The Leadership Program's Violence Prevention Project" OR "VPP" OR "Too Good for Violence" OR "Tools for Getting Along Together" OR "Tools of the Mind" OR "Transformative Life Skills and Dynamic Mindfulness" OR "WITS" OR "Zippy's Friends" **AND**

4. [Abstract and title code] (adapt* OR modif* OR alter* OR adjust* OR attun* OR chang*) [proximity operator of 5] (language* OR race* OR racial* OR ethnic* OR culture* OR cultural*)

Third, for studies that ultimately ended up meeting inclusion criteria, a manual ancestral search was conducted of the references of articles that met all criteria to determine if there were any of the search terms used in the same combination as the original search

Inclusion and Exclusion Criteria

Following the electronic search process, an initial screening of titles and abstracts was conducted, followed by a full text review. Inclusion and exclusion criteria were established to identify studies that would be suitable for the scope of this review. The

studies eligible for inclusion in this review were: (a) peer-reviewed empirical studies or dissertations; (b) conducted in PreK-12 school settings in the United States; (c) available in English; and (d) included a cultural adaptation to a SEL program. A cultural adaptation was defined as, "the systematic modification of an evidence-based treatment (EBT) or intervention protocol to consider language, culture, and context in such a way that [was] compatible with the client's cultural patterns, meanings, and values" (Bernal et al., 2009, p. 362). SEL programs were defined as including the following elements for the purpose of this review: (1) included lessons and/or activities for any grades between PreK-12; (2) had a primary focus on SEL; (3) had well-defined activities that is aligned with the practice of SEL; and (4) had well-defined activities that support one or more of the CASEL SEL competencies: self-awareness, self-management, social awareness, relationship skills, responsible decision-making.

The studies that were excluded contained the following characteristics: (1) studies whose main focus was on a broader system (e.g. MTSS, PBIS, etc.) that subsumed SEL programming; and/or (2) studies not conducted in the United States; and/or (3) studies without a full length article available and/or; (4) the only adaptation made to the SEL program was adapted for a reason other than meeting cultural needs of the participants (e.g. mode of delivery; telehealth, developmental, etc.).

Coding

To better understand the types, purposes, and associations to student outcomes of cultural adaptations to SEL programs, the following information was extracted from each eligible study: author(s), year of publication, research question(s), study design, type of SEL program used (e.g. small-group, class-wide), number of lessons and/or activities implemented, name of SEL program (if there was one), length of program (e.g. weeks, days, years), number of participants in the study, implementer of the SEL program, CASEL

Competency targeted by the SEL program, demographic characteristics of the participants and the implementer (e.g. grade, race/ethnicity, gender/sex, years of experience, role, etc.), types of cultural adaptations being made to the SEL program, purposes of the cultural adaptations being made to SEL program, student outcomes associated with cultural adaptations to the SEL programs (i.e., measures used, dependent variables: primary and secondary, effect sizes, etc.), social validity (e.g. measures of accessibility, feasibility, etc.), treatment integrity, and quality appraisal score. See Codebook in Appendix A.

Quality Appraisal Tool

The primary author reviewed selected studies and evaluated them with the Quality Assessment with Diverse Studies (QuADS; Harrison et al., 2021) appraisal tool. This tool evaluates psychological research beyond experimental design and was thus selected as studies reviewed could potentially use a variety of research methods. QuADS was developed to address the limitations of the Quality Assessment Tool for Studies with Diverse Designs (QATSDD; Sirriyeh et al., 2011). A limitation of the QATSDD included a lack of clarity surrounding the description of assessment items (Fenton et al., 2015). The QuADS now provides more in-depth descriptions for each item along with a high inter-rater reliability and face and content validity evidence to support use (Harrison et al., 2021). This tool is discussion-based; therefore, a second rater examined each of the selected articles. Both raters met and consensus built on the quality of each study. The results of the quality appraisal are in Table 5. A copy of the directions for use and the QuADS tool can be found in Appendix B and Appendix C, respectively.

Data Analysis Plan

For the first research question, “What types of cultural adaptations have been made to SEL programs in the empirical literature?” the researchers categorized the cultural adaptations into surface and deep structure features based on the cultural sensitivity model

(Resnicow et al., 1999). These types were more specifically categorized into common features of surface (e.g. translation of materials) and deep (e.g. included cultural values in intervention design) structural adaptations that were compiled by Barrera et al. (2013) from Mier et al. (2010) and Kreuter and Wray (2003). See Codebook in Appendix A for further information. The second research question, "What are the stated purposes of the cultural adaptations made to SEL programs in the empirical literature?" was answered by extracting the author(s)' stated purposes for making the cultural adaptation to the SEL program. The authors of the present study reviewed the stated purposes, summarized them, and grouped similar purposes together in the analysis. If no purpose was given, a "missing" code was assigned (see Codebook in Appendix A).

For Research Question 3, "To what extent are these cultural adaptations of SEL programs associated with student outcomes?" the present systematic review relied on the reported student outcomes associated with cultural adaptations to the SEL programs and reported social validity. One type of reported outcome was an effect size. Authors in the included studies reported Cohen's *d* Cohen, 1988), partial eta squared (η_p^2), and phi (φ). If studies did not report an effect size and Cohen's *d* could be calculated, the authors of this review calculated the Cohen's *d* effect size. When effect sizes could not be calculated and no other effect size type were provided, other descriptive statistics or qualitative data were provided. Effect sizes for non-adapted programs (e.g. control and non-adapted version conditions) were not provided.

In addition, the authors examined how those involved with the studies perceived the social validity of the SEL programs' cultural adaptations. Social validity information was extracted using quantitative and/or qualitative data related to social validity measures. When ambiguous information was provided in text, study authors were emailed for specific

details only if extractable content was present but unclear to both the primary researcher and another graduate student or university faculty member.

Results

Search Results

A total of 8,947 articles were found in the electronic search process: ERIC (*k* = 666), Education Source (*k* = 681), PsycInfo (*k* = 1822), Medline (*k* = 3,493), and Dissertations and Theses Global (*k* = 2,285).

After removing duplicates, a total of 7,190 articles remained. First, titles and abstracts were screened for relevance, after which 82 studies remained. Second, a full text review was conducted, removing 71 articles for not meeting the following inclusion criteria: not a peer reviewed study or dissertation (*k* = 4), not conducted in a PreK-12 school setting in the United states (*k* = 17), not available in English (*k* = 0), did not include a cultural adaptation to an SEL program (*k* = 12), the study did not contain any PreK-12 SEL lessons (*k* =33), study whose main focus was on a broader system (*k* = 4), studies without a full length article available (*k* = 0), and the adaptation made to the SEL program was adapted for a reason other than meeting the cultural needs of the participants (*k* = 12). Finally, one dissertation (Castro Olivo, 2007) was removed to avoid duplication as it was subsequently published in a peer-reviewed journal (Castro-Olivo & Merrell, 2012) which was the article used for analysis in this review. The ancestral search yielded no additional studies for analysis. The final review included 11 studies comprised of peer-reviewed studies (*k* = 8, 73%; Castro-Olivo & Merrell, 2012; Castro-Olivo, 2014; Castro-Olivo et al., 2022; Cramer & Castro-Olivo, 2015; Elek et al., 2010; Graves et al., 2017; Hecht et al., 2018; Moran et al., 2019) and dissertations (*k* = 3, 27%; Hopson, 2008; Kurtz, 2020; Martinez, 2011). See Figure 1 for a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram outlining the search and retrieval process.

Study Characteristics

Table 2 presents the descriptive characteristics of the studies. Information regarding the setting of the school(s) and/or district(s) along with the analysis type and method were provided. For studies that provided the urbanicity of their study location (*k* = 6, 55%), three (50%; Graves et al., 2017; Hecht et al., 2018; Hopson, 2008) were conducted in urban settings, one (17%; Cramer & Castro-Olivo, 2015) was conducted in a suburban setting, and two (33%; Hecht et al., 2018; Moran et al., 2019) were conducted in a rural setting. All studies (*k* = 11, 91%; Castro-Olivo & Merrell, 2012; Castro-Olivo, 2014; Castro-Olivo et al., 2022; Cramer & Castro-Olivo, 2015; Elek et al., 2010; Graves et al., 2017; Hecht et al., 2018; Hopson, 2008; Kurtz, 2020; Martinez, 2011; Moran et al., 2019) provided a state or region in which their study was conducted. For studies that provided either, seven (70%; Castro-Olivo & Merrell, 2012; Castro-Olivo, 2014; Castro-Olivo et al., 2022; Cramer & Castro-Olivo, 2015; Elek et al., 2010; Martinez, 2011; Moran et al., 2019) were conducted in the western U.S.; one (10%; Hecht et al., 2018) was conducted in the Midwest; one (10%; Kurtz, 2020) was conducted in the eastern U.S., and one (10%; Graves et al., 2017) was conducted in the Mid-Atlantic region. Seven of the studies (64%; Castro-Olivo, 2014; Castro-Olivo et al., 2022; Elek et al., 2010; Graves et al., 2017; Hecht et al., 2018; Hopson, 2008, Kurtz, 2020) conducted between group analyses and four of studies (36%; Castro-Olivo & Merrell, 2012; Cramer & Castro-Olivo, 2015; Martinez, 2011; Moran et al., 2019) conducted within group analyses.

Demographic information regarding participants and those who implemented the SEL programs can be found in Table 3 and Table 4, respectively. The number of students involved in the studies ranged from 20 to 2,781 in total, culturally adapted SEL programs were implemented with 5,173 students across studies. Three studies (27%; Graves et al., 2017; Kurtz, 2020; Moran et al., 2018) included students who were primarily in elementary

grades (ie. Kindergarten through 5th grade). Three studies (27%; Elek et al., 2010; Hecht et al., 2018; Martinez, 2011) included students who were primarily in middle school grades (ie. 6th through 8th grade). Three studies (27%; Castro-Olivo & Merrell, 2012, and Cramer & Castro-Olivo, 2015; Hopson, 2008) included students in primarily high school grades (ie. 9th through 12th grade). Two studies (18%; Castro-Olivo, 2014; Castro-Olivo et al., 2022) included students in middle and high school. Most studies (*k* = 7, 64%; Castro-Olivo & Merrell, 2012; Castro-Olivo, 2014; Castro-Olivo et al., 2022; Cramer & Castro-Olivo, 2015; Elek et al., 2010; Kurtz, 2020; Martinez, 2011) had predominantly Mexican/Mexican-American, Latino/Latine, or Hispanic student populations. One study (9%; Graves et al., 2017) had a focus on African American students. One study (9%; Hecht et al., 2018) had a focus specifically on European Americans. Two studies (18%; Hopson, 2008; Moran et al., 2019) had a predominantly equal divide between the number of Hispanic and white students they included. In terms of gender or sex assigned at birth, most studies (*k* = 10, 91%; Castro-Olivo & Merrell, 2012; Castro-Olivo, 2014; Castro-Olivo et al., 2022; Cramer & Castro-Olivo, 2015; Elek et al., 2010; Hecht et al., 2018; Hopson, 2008; Kurtz, 2020; Martinez, 2011; Moran et al., 2019) included both girls and boys or females and males in their study. Of these ten studies, three studies reported data where students could select a "other" or "choose not to report '' option. One study (9%; Graves et al., 2017) strictly included male students.

Implementer data (i.e., characteristics of the interventionist who implemented the SEL programs) were widely not reported. The number of implementers for each study ranged from two to 32. Two studies (27%; Elek et al., 2010; Moran et al., 2019) did not report the number of implementers. Most implementers were teachers or school counselors (*k* = 6, 55%; Castro-Olivo & Merrell, 2012; Castro-Olivo, 2014; Castro-Olivo et al., 2022; Hecht et al., 2018; Hopson, 2008; Kurtz, 2020) and four studies (36%; Cramer & Castro-

Olivo, 2015; Graves et al., 2017; Martinez, 2011; Moran et al., 2019) utilized graduate students as implementers. Only one study (9%; Kurtz, 2020) provided the years of experience their teachers had, ranging from one to 19 years. Seven studies (64%; Castro-Olivo & Merrell, 2012; Castro-Olivo, 2014; Castro-Olivo et al., 2022; Cramer & Castro-Olivo, 2015; Graves et al., 2017; Hecht et al., 2018; Hopson, 2008) stated teachers received training for the implemented SEL program. One study (9%; Kurtz, 2020) stated that teachers had no experience with the SEL program, and three studies (27%; Elek et al., 2010; Martinez, 2011; Moran et al., 2019) did not provide this information. Only three studies (27%) noted the race/ethnicity of their implementers. Two studies (18%; Castro-Olivo, 2014; Castro-Olivo et al., 2022) had 100% Latina women, but it is important to note that these studies were completed by the same author on the same sample of participants. One study (9%; Kurtz, 2020) had predominantly white with a few Black, African American, or Hispanic implementers. The other studies did not present this information (k = 8, 73%). Even less demographic information was provided for the gender or sex assigned at birth of the implementers. Only two studies (18%; Castro-Olivo, 2014; Castro-Olivo et al., 2022) reported the implementers were 100% women, which again, one study was a secondary analysis of the primary study. The other studies did not report this information.

Quality Appraisal

The minimum and maximum total score possible was zero and 39, respectively. Each study was rated on a zero to three Likert-type scale where, in general, zero represented no mention of the quality indicator, and three represented completeness and detail of the quality indicator. On average, articles scored 87% (low = 61.5%; high = 97.4%) of total possible points. Categories with the lowest average score were “Justification for analytic method selection” (average score = 1.6), “Theoretical or conceptual underpinning to the research” (average score = 2.2), and “Recruitment data provided” (average score =

2.3). Categories where all included studies obtained the highest score of a three included the "The study design is appropriate to address the stated research aim(s)," "Appropriate sampling to address the research aim/s," and "Evidence that the research stakeholders have been considered in research design or conduct." The full quality appraisal can be found in Table 5.

SEL Program Characteristics

The culturally adapted programs can be found in Table 6. The most common SEL programs that were culturally adapted were versions of Merrell's Strong Start, Strong Kids, and Strong Teens interventions (Merrell et al., 2007a; Merrell et al., 2007b; Merrell et al., 2007c) (k = 6, 55%). Three studies (27%; Elek et al., 2010; Hecht et al., 2018; Hopson, 2006) adapted the SEL program keepin' it REAL (Gosin et al., 2003), a drug use prevention program that also targets SEL skills. One program (9%; Moran et al., 2019) adapted the Stealing the show, Turning it over, Accompanying others, and Coaching compassion (STAC; Midgett et al., 2015) bullying prevention SEL program. One study (Elek et al., 2010) implemented the adapted SEL programs for school-wide use, five (45%; Castro-Olivo & Merrell, 2012; Castro-Olivo, 2014; Castro-Olivo et al., 2022; Hecht et al., 2018; Hopson, 2008) for class-wide use, three (27%; Cramer & Castro-Olivo, 2015; Graves et al., 2017; Martinez, 2011; Moran et al., 2019) for small group use, and one study (9%; Kurtz, 2020) examined the adapted SEL program for class-wide and small group use. The number of SEL lessons per program ranged from three to 14 lessons, with most programs completing one lesson per week. Two studies (18%; Castro-Olivo et al., 2022; Cramer & Castro-Olivo, 2015) targeted all five CASEL competencies explicitly while seven studies (64%; Castro-Olivo & Merrell, 2012; Castro-Olivo, 2014; Graves et al., 2017; Hopson, 2008; Kurtz, 2020; Martinez, 2011; Moran et al., 2019) targeted all 5 CASEL competencies implicitly. See Table 6 for examples of how implicitly targeted competencies were coded to fit the CASEL

competencies. One study (9%; Elek et al., 2010) targeted four of the five CASEL competencies implicitly with one additional non-CASEL related competency, and one study (9%; Hecht et al., 2018) targeted only one of the five CASEL competencies implicitly.

Treatment Integrity and Social Validity

Studies used different approaches to measure treatment integrity, including interviews and observations conducted by researchers along with audiotapes and checklists completed by implementers. Five studies (45%; Castro-Olivo & Merrell, 2012; Elek et al., 2010; Graves et al., 2017; Hopson, 2008) included outcomes of their treatment integrity with two providing strictly quantitative results (Graves et al., 2017; Kurtz, 2020), two reporting strictly qualitative results (Elek et al., 2010; Hopson 2008), and one (Castro-Olivo & Merrell, 2012) detailing both quantitative and qualitative results. In general, studies reported high treatment integrity. See Table 7 for details.

A total of eight studies (73%; Castro-Olivo & Merrell, 2012; Castro-Olivo, 2014; Castro-Olivo et al., 2022; Cramer & Castro-Olivo, 2015; Graves et al., 2017; Kurtz, 2020; Martinez, 2011) provided measures of social validity. However, only seven of the eight studies (88%; Castro-Olivo & Merrell, 2012; Castro-Olivo, 2014; Cramer & Castro-Olivo, 2015; Graves et al., 2017; Kurtz, 2020; Martinez, 2011) provided social validity outcome data (see Table 10). Data came in the form of student interviews (k = 2, 23%; Martinez, 2011 and Moran et al., 2019), teacher interview (k = 1, 14%; Graves et al., 2017), student rating scales (k = 4, 57%; Castro-Olivo & Merrell, 2012; Castro-Olivio, 2014; Cramer & Castro-Olivo, 2015; Martinez, 2011), teacher rating scales (k = 3, 43%; Castro-Olivo & Merrell, 2012; Graves et al., 2017; Kurtz, 2020). Student interviews, student rating scales, and teacher rating scales all had high levels of acceptability, feasibility, and satisfaction. Students felt they had learned SEL skills and felt more confident to use skills. Teachers felt that they observed students using skills and believed the interventions were feasible

overall. One study (14%; Graves et al., 2017) reported that teachers desired a stronger focus on Black boys, resilience, and more time.

Research Question 1: Types of Cultural Adaptations

Nine studies (82%; Castro-Olivo & Merrell, 2012; Castro-Olivo, 2014; Castro-Olivo et al., 2022; Cramer & Castro-Olivo, 2015; Graves et al., 2017; Hopson, 2008; Kurtz, 2020; Martinez, 2011; Moran et al., 2019) leveraged both surface and deep level cultural adaptations. Two studies (18%; Elek et al., 2010; Hecht et al., 2018) used only deep level adaptations. Surface adaptations included bilingual and bicultural implementers, translation of materials, inclusion of ethnic lifestyle elements (e.g. using home-language to describe feelings), inclusion of same race/ethnicity role models (e.g. selected books where main character is a child of color), and other surface level adaptations (e.g. changed lesson length to meet alternative school length, options for culturally appropriate language). Deep level adaptations included incorporating cultural values in SEL program design (e.g. collaborated with rural teachers to incorporate rural culture in the programs), incorporating cultural values in SEL program implementation (e.g. taught skills reflecting cultural values such as respeto), involving family in the SEL program (e.g. formal graduation party held for students and their families), adjusting materials to the literacy level of the participant (e.g. incorporated activities with more visuals to support English language learners), and other deep level adaptations (e.g. included metaphors and folktales, taught about trauma responses, acculturative stress, ethnic pride, and involved implementers in advocacy efforts). See Table 8 and Appendix D (Kurtz, 2020) for a complete list of cultural adaptations.

Research Question 2: Stated Purposes of Adaptations

All researchers provided one or more stated purposes for their cultural adaptations. Eleven studies (100%; Castro-Olivo & Merrell, 2012; Castro-Olivo, 2014; Castro-Olivo et al.,

2022; Cramer & Castro-Olivo, 2015; Elek et al., 2010; Graves et al., 2017; Hecht et al., 2018; Hopson, 2008; Kurtz, 2020; Martinez, 2011; Moran et al., 2019) mentioned adapting the programs to fit the population and/or setting (e.g. ELL, CDL, racial/ethnic group, rural, urban, alternative school). See Table 8 for more information. More restricted purposes included meeting the teachers expectations for what SEL skills should be taught (9%; Graves et al., 2017), adapting the program to fit an older age group in an alternative school setting (9%; Hopson, 2008), addressing the needs of students experiencing acculturation (9%; Elek et al., 2010), supporting the high needs and limited resources of the schools (9%; Kurtz, 2020), and having a focus group to guide the adaptation process (9%; Kurtz, 2020).

Research Question 3: Student Outcomes

Studies examined a variety of different outcomes including, but not limited to, drug use, intentions to accept drugs (e.g. alcohol and marijuana), drug resistance strategies, anti-drug norms, educational aspirations, psychological adjustment, anger management, self-esteem, ethnic pride, gang involvement, SEL knowledge, internalizing symptoms, externalizing symptoms, sense of belonging, internalizing symptoms, and student resiliency. See Table 9 for all the measured student outcomes. There were three (27%; Elek et al., 2010; Hecht et al., 2018; Hopson, 2008) SEL programs that had a focus on drug prevention, using the *keepin'* it REAL program. For the drug prevention SEL studies, student outcomes focused on alcohol, marijuana, cigarette, and tobacco use, intentions to accept substances, refusal efficacy, and substance use norms. The other SEL programs focused on student knowledge of SEL, internalizing symptoms, externalizing symptoms, sense of belongingness, acculturative stress levels, and gang involvement. See Table 9 for details.

In general, the drug prevention-focused SEL programs found mixed effects. In comparison to control groups, only one of the three drug prevention SEL programs (33%; Hecht et al., 2018) found more improvement in terms of decreased substance use for

students in the adapted control condition than in the control conditions. The other two drug prevention-focused SEL studies found mostly non-statistically significant differences between groups but change in the desired direction (Hopson, 2006) or mostly non-statistically significant differences between groups with change in the undesired direction (Elek et al., 2010).

The other eight SEL programs found a range of effects regarding the adapted SEL programs. Of the eight studies, five (63%; Castro-Olivo & Merrell, 2012; Castro-Olivo et al., 2022; Cramer & Castro-Olivo, 2015; Martinez, 2011; Moran et al., 2019) found mostly non-statistically significant mean differences or changes between control and intervention groups or from pre to post assessment. Three of the eight studies (37%; Castro-Olivo, 2014; Graves et al., 2017; Kurtz, 2020) found mostly small to large, positive (i.e., change in the desired direction) effects of the culturally adapted interventions. Four studies (50%; Castro-Olivo & Merrell, 2012; Castro-Olivo, 2014; Graves et al., 2017; Moran et al., 2019) measured changes in SEL knowledge, and all four studies (100%) found the adapted SEL program greatly increased student knowledge of SEL skills. Of the four studies (50%) that measured internalizing behaviors, two of the studies (Castro-Olivo & Merrell, 2012; Cramer & Castro-Olivo, 2015) found no statistically significant difference between groups, and two of the studies (Castro-Olivo, 2014; Kurtz, 2020) found small to large effect sizes of the culturally adapted program on improving internalizing behaviors. Only one study (13%; Graves et al., 2017) examined externalizing behaviors and found a small effect of the culturally adapted program on decreasing externalizing behaviors. See Table 9 for more specific student outcomes.

Discussion

Culturally relevant practices for students, including the adaptation of extant interventions, has become more prevalent as the diversity of needs in the U.S. grows

(Cooper et al., 2020). Systematic reviews and meta-analyses have been conducted regarding cultural adaptations to parent interventions (Baumann et al. 2015) and psychological interventions (Chowdhary et al., 2014; Nagayama Hall et al., 2016); however, none to our knowledge have been conducted regarding cultural adaptations to social-emotional learning programs—despite the breadth of cultural adaptations that have been conducted domestically (e.g. Castro-Olivo et al., 2022; Moran et al., 2019) and internationally (Ferrer-Wreder et al., 2021; Marsiglia et al., 2022; Matsuba et al., 2021). This study is the first to systematically review the types of cultural adaptations made to SEL programs, the purposes of the adaptions, and the student outcomes associated with the cultural adaptations to SEL programs conducted in the U.S. This systematic review captured how studies leveraged both surface and deep-level cultural adaptations to their SEL interventions for the purpose of fitting the specific racial/ethnic and geographical locations of their students. Cultural adaptations to SEL programs were associated with increased student knowledge regarding SEL skills and positive social validity evidence. However, effects on other student outcomes, such as less drug use and internalizing behaviors had mixed findings.

In reflecting on the contributions of this work to knowledge, it is important to consider the representativeness of the population of students and implementers included in this review. The student population represented in this review was predominantly Latine. Given the diversity of the U.S. population, there remains much work to be done to better understand how cultural adaptations might be applied to support other populations. Yet, when examining and making cultural adaptations, a simplistic focal point can be the race or ethnicity of a population. Indeed, race and ethnicity were discussed in every study reviewed. However, almost every study contained a more nuanced view of their populations.

One example of this view was the examination of the locale of the students. Fiske (2002) briefly discusses modifications to a physical environment as part of a culture; however, other studies, such as those from Hecht et al. (2018), Hopson (2006), and Moran et al. (2019), underscore the importance of examining the location of different racial and ethnic groups when making cultural adaptations because of the variability of beliefs, customs, and practices within a racial or ethnic group. These studies went beyond geographical region as a part of culture (e.g. state) and examined factors as broad as urbanicity (e.g. schools in rural areas) to as narrow as school-setting type (e.g. alternative school settings) in their definitions of culture. These considerations the authors of the reviewed studies took when examining the construct of culture demonstrated there are several intricacies of identity to consider when creating cultural adaptations. However, this intricate view of identity was glaringly absent for implementers of the SEL programs. With only three studies (27%) providing the race or ethnicity of their SEL program implementers, there appears to be a fracture in our knowledge of how surface-level adaptations to who is presenting the SEL program could influence student outcomes. As researchers consider exploring the impact of racial/ethnic mismatch between teachers and students, this is a critical area of study.

This systematic review revealed SEL researchers made more than simple adaptations to SEL programs, much like other systematic reviews examining cultural adaptations (Alcántara et al., 2021; Marshall et al., 2021; Vincze et al., 2021). Captured in this systematic review, most studies made both surface and deep-level cultural adaptations to their SEL interventions. This is beneficial as a combination of both may sometimes be necessary when individuals and their interventions have a cultural mismatch (Resnicow et al., 1999). However, having both adaptation types in an intervention unexpectedly had varying results. Although deep-level adaptations attempt to delve into the ecological factors

that may influence how individuals could better respond to an intervention, both surface and deep-level adaptations have been found to have varying results when introduced individually or together. For example, a systematic review that examined cultural adaptations in health-related areas found that the level of adaptation did not influence the effectiveness of the intervention (Nierkens et al., 2013). This finding is similar to the findings of this systematic review where the negative, neutral, and positive student outcomes did not necessarily stem from a particular level of adaptation. The extent of the effectiveness of these cultural adaptations is limited by the strikingly few numbers of SEL program types that were adapted.

When considering the Merrell suite (ie. Strong Start, Strong Kids, and Strong Teens) and keepin' it REAL suite (e.g. keepin' it REAL-rural, keepin' it REAL-multicultural, etc) of SEL programs as unique program types, there were only three unique programs included—limiting the field's knowledge of what types of SEL programs can be effectively and feasibility adapted. Of the extant data, the number of adaptations made to each study were expected. With regard to the few numbers of SEL program types that were adapted, more consideration is needed for the environmental factors that could influence student outcomes. For example, in the drug-prevention SEL programs, several reasons that go beyond a culturally adapted SEL program could contribute to student drug use such as family and community norms, social pressures, and curiosity (Drug Enforcement Administration; DEA, 2023). However, in an effort to fully capture the state of the literature, the authors thought it was important to include these studies as well, particularly given their inclusion in the CASEL list of SEL programs.

Another consideration as to why there were mixed results could be the types of measures used in the included studies. Indeed, a recent meta-analysis of 111 SEL-focused studies found that many SEL measures evaluate outcomes broadly (e.g. depression, drug

use) rather than specifically (e.g. identifying emotions, problem solving) which makes following growth in SEL skills and knowledge over time difficult (Ura et al., 2020). Reinforcing this finding, the outcomes measured in many of these programs were broadly defined (e.g. Hopson (2006) examined drug use; Martinez (2011) examined educational aspirations; Cramer and Castro-Olivo (2015) examined internalizing symptoms). To better understand how cultural adaptations impact diverse student outcomes, more sensitive measures that directly examine SEL skills could be used.

Further, there could be more to consider from a foundational perspective of CRP. In this review, there was a grounding assumption that cultural adaptations can be beneficial supports to meeting goals in several areas of life, including in education (Luevano & Collins, 2020); yet benefits were not evident across all outcomes. CRP asserts the importance of relatable education. Despite this knowledge, the student outcomes related to the cultural adaptations conducted in the reviewed studies were mixed. Ladson-Billings' (1995) conceptualization of CRP requires three components: student learning, cultural competence, and *critical consciousness*. Originally conceptualized by Paulo Freire in 1970, critical consciousness is "learning to perceive social, political, and economic contradictions, and to take action against the oppressive elements of reality" (Freire, p. 35). More directly, critical consciousness is the ability to understand the roots of social inequities (California Department of Education, 2023). Although the integration of student learning could be seen in the measurements of student SEL knowledge in a few studies, and cultural competence was integrated into the creation of the cultural adaptations, critical consciousness seemed to be missing in the creation of the cultural adaptations and in the outcomes measured. Thus, when grounding the importance of cultural adaptations in SEL programs on the theory of CRP, researchers could examine how the cornerstone aspect of critical

consciousness could be infused into the process to support increased, positive student outcomes.

Lastly, the strong, social validity evidence of cultural adaptations found by the social validity measures echoed the findings of other systematic reviews regarding cultural adaptations (e.g. Sivaraman & Fahmie, 2020). Despite only some of the studies finding positive impacts of the cultural adaptations on supporting decreased substance use, increased SEL knowledge, and decreased internalizing and externalizing behaviors, almost all studies in this review found the adapted interventions to be acceptable, feasible, and satisfactory by students and teachers alike. Given these strong social validity findings, research should continue in this area despite the mixed student outcomes found in this single review.

Limitations

Although the quality of these studies may be considered relatively high, with an average of 88% of appraisal items met, it is important to note that not all quality indicators are equal, and thus a more nuanced understanding of study quality is needed. Additionally, there were several other limitations noted by the authors of the studies in this systematic review. First, authors in this review noted attrition, sample size, and fidelity (Cramer & Castro-Olivo, 2015; Graves et al., 2017; Hopson, 2006), dated curriculum (Elek et al., 2010), heterogeneity within the population (Hecht et al., 2018), use of self-report measures (Martinez, 2011), lack of true control groups (Castro-Olivo, 2012; Kurtz, 2020; Moran et al., 2020), choice of measurement methods (Castro-Olivo, 2014; Castro-Olivo et al., 2022) as limitations. Second, no researcher conducted adaptations individually, all adaptations were made in conjunction with multiple other adaptations, presenting confounds to understanding unique influences of adaptations. Moreover, two of the studies drew

conclusions based on an identical sample of participants, which introduces redundancy in terms of the influence of these studies on the overall conclusions drawn.

Beyond the scope of the limitations of each individual study, this review had several limitations that should be considered. The first being that the definition of SEL, culture, and culturally adapted are vast concepts that are difficult to define. The numerous definitions and variations of the terms make capturing all the literature potentially related to the two concepts an onerous task. For example, in reviewing the literature, the terms "tailored" and "cultural grounding" (instead of adapted) appeared. However, these terms were not included in the original search, so it is possible that other relevant studies were not included in this review. Further, programs that were developed for specific populations (e.g. Graves & Aston., 2018) and their efficacy were not included, thus results from these studies were not considered in this review. Another limitation included the selected quality appraisal tool. One of the items in this tool included a specification of the city a study took place for the study to receive full credit. Given the sensitivity of educational research, this level of specificity is often excluded to preserve anonymity. Therefore, the breadth of the quality appraisal tool that is helpful to evaluate multiple research designs did not necessarily account for the anonymity needed in education research. Lastly, efforts were made to contact study authors when ambiguous information was provided. Three authors were emailed during the coding process, and thus codes were assigned based on information provided. The accuracy of this information, and potential for response biases should be considered in this context.

Implications for Future Research

This study and other systematic reviews of cultural adaptations have discussed limitations in terms of few studies in this area for generalizability (Al Maskari et al., 2018). Future studies could examine what types of SEL programs are most suitable for adaptation.

Moreover, given the few groups receiving adapted programs, more research could be conducted with other populations that were clearly missing or under-represented within this study (e.g. Indigenous populations, Asian/Pacific Islander populations, Black and African American, rural and suburban populations, etc.). It is important to recognize that even within these populations that were not or under-studied, there is great diversity within these racial and ethnic categories due to geographic location and specific cultural practices. Thus, future research should take care to specify which populations they are adapting interventions for, where these populations are located, and name specific details regarding racial and ethnic identities as articulated by the studied population. Finally, researchers could further explore best practices in culturally adapting interventions to determine and maximize the benefits of deep and surface level adaptations.

Implications for Practice

Cultural relevance of programs is necessary for SEL to move beyond the benefits it reaps through positive school climates and increased academic achievement. Cultural relevance is indeed necessary for *Transformative* SEL, a form of SEL that pushes students to understand and redistribute power, to engage civically in their communities, and recognize the importance of identity, agency, belonging, and collaborative problem solving (Jagers et al., 2021). Indeed, many of the studies included in this review collaborated with individuals from the community and students impacted by the adaptations to ensure cultural adaptations were beneficial and not harmful. Another consideration for practice is the amount of time, money, and expertise available to adapt typically "universal" SEL programs for specific populations. Given the few studies conducted in this area, districts, schools, and teachers will likely need to weigh the perceived benefits and costs of adapting interventions. However, the feasibility presented by this review is promising to move toward interventions that may be beneficial to all students. However, Ogbu (1987) pushes

for understanding that student success cannot be solely attributed to the relatability of curriculum, but also the relationship between oppressive and oppressed classes of individuals. Thus, districts, schools, and teachers could consider if cultural adaptations are being used as a temporary solution to larger, systemic issues in their district.

Conclusion

For all students to benefit from SEL programs, educators, researchers, and communities must consider the fit between programs and recipients. Minimal work has been done to understand why cultural adaptations are important, how cultural adaptations can be made, and the student outcomes associated with cultural adaptations. This systematic review has demonstrated that few cultural adaptations have been made to SEL programs. The lack of research in this area calls for future studies to be conducted to better understand in what ways adaptations can be made to serve diverse populations.

Chapter 3

Study 2: A Preliminary Investigation of the Be Good People Social Emotional Learning Program

Abstract

Social and emotional needs among children are rising in the United States, and educators are seeking effective ways to support their students. One recently developed open-access social emotional learning (SEL) program, titled Be Good People (BGP), has nearly 2,000 registered users (either schools or school districts) across the world, but no research to support its use. An initial feasibility study was conducted in a rural Midwestern elementary school to understand how the BGP program was associated with students' social, emotional, and academic outcomes, along with how teachers perceived the acceptability, practicality, and cultural relevance of the program. A convenience sample of seven teachers from first (*n* = 3), second (*n* = 1), third (*n* = 2), and fourth (*n* = 1) grade classrooms provided quantitative and qualitative information regarding their experience implementing BGP in their classrooms across a school year. De-identified student level social, emotional, and academic data were provided from 111 students across participating classrooms, and teachers were asked to share their perceptions of BGP through interviews and rating scales. Results indicated there was some improvement from Time 1 to Time 2 of this study in terms of students' social skills and academic performance. There were no statistically significant changes in students' social, emotional, and behavioral risk ratings. Teachers rated and described BGP as practical, but there were concerns for acceptability, including the program not being developmentally appropriate for young elementary students. Additionally, BGP requires more work in the area of cultural relevancy. Implications for practice and research are discussed.

Keywords. Social emotional learning, Be Good People, cultural relevance

A Preliminary Investigation of the Be Good People Social Emotional Learning Program

Children and adolescents are growing up in a more complex world, seemingly, each year. In this century, with immediate access to the news at their fingertips, students in the United States have witnessed a pandemic, uprisings for racial equity, wars, and political strife. Indeed, recent data indicates that students are struggling in numerous ways. Educators, parents, and caregivers have noticed an uptick in problem behaviors, greater internalizing behaviors, and fewer social skills among students (NCES, 2022; Watson et al., 2022). Given that students spend the majority of their time in schools, it is crucial that schools serve as an environment where students can continue to develop their social and emotional skills. In recent years, there has been a growing push to formally teach these skills through evidence-based social emotional learning (SEL) programs. These programs aim to support students in managing emotions, making healthy choices, developing their identities, creating and keeping strong relationships (Collaborative for Academic and Social Emotional Learning; CASEL, n.d.-a).

However, commercially available evidence-based SEL programs can be extremely costly (Buchanan et al., 2009). Thus, some school districts have turned to creating their own SEL programs to meet the needs of students in their district. These locally developed programs can be cost-effective for school districts (Bailey et al., 2019); yet there is not necessarily a strong research base for these programs. One such program is the Be Good People (BGP) program, which is widely available and accessible (see link to the webpage here); however, it lacks evidence behind its effectiveness in schools. As a result, uncertainties about its implementation and effectiveness contribute to a limited understanding of the program's merits and impacts. As such, the purpose of this investigation was to conduct an initial, feasibility study of the BGP SEL program to gather

preliminary data regarding: (1) fidelity of implementation, (2) students' social-emotional, academic and behavioral outcomes, and (3) teacher perceptions of the practicality, acceptability, and cultural relevance of the program.

Social Emotional Learning (SEL)

SEL as concept began rising to popularity in the 1990s (CASEL, n.d.-a). The need for public education to move beyond pure academics and to include SEL has become increasingly clear as more research establishes that SEL boosts student achievement and overall school climate (Kendziora & Yoder, 2016). The highly influential meta-analysis conducted by Durlak et al. (2011) of over 270,000 kindergarten to high school students indicated that students who have increased social and emotional skills have higher academic achievement and stronger relationships with peers. Students are also able to better handle difficult situations, unpredictable moments, and change (Durlak et al., 2011). These results were replicated with a more recent meta-analysis with studies from across the world, including over 575,000 students (Cipriano et al., 2023). The researchers found that students who participated in SEL interventions showed growth in areas such as school climate/safety, civic attitudes/behaviors, SEL skills, peer relationships, attitudes/beliefs, and prosocial behaviors compared to students not involved in SEL interventions. Although these findings provide critical support for the overall impact of SEL programs, not all programs are created equal. Within the landscape of SEL, there are many programs educators can choose to implement to teach these essential skills.

The landscape of SEL programs is broad; CASEL alone has 85 listed programs on their website (CASEL, 2023). These programs can be filtered into many categories, including but not limited to the following: grade bands, program approaches, implementation support offerings, settings in which SEL can be promoted across (e.g. class-wide, school-wide, community-wide or at home), school urbanicity, and student characteristics (e.g. race,

income level)—demonstrating the variety of SEL programs available (CASEL, 2023). Further, CASEL continues to accept applications for SEL programs to be designated as CASEL-approved. Moreover, Cipriano et al. (2023) identified 252 unique programs in their meta-analysis, where 210 were named and 42 were unnamed. Of the unnamed programs, some authors described creating their own SEL programs as a form of cost-effectiveness for the schools in which they were implemented. Examples included an unnamed, six-week mindfulness program (Britton et al., 2014) and an unnamed, eight-week mindfulness program (Johnstone et al., 2020). Another school district created a two-week SEL program that took its foundation from Alaskan SEL standards and concepts from the Six Seconds (Six Seconds, n.d.) framework (Martinez, 2016). Specifically, these programs were all universal (i.e., school wide) SEL programs which are designed to teach all students SEL skills.

Universal SEL programs

Universal SEL programs are used in many schools across the country as a method to teach all students social and emotional skills. As opposed to targeted or intensive programs which are designed to meet the specific needs of students, universal programs are designed to be "applicable to all." Driven in part by a growing number of states adopting SEL standards (currently, 27 states have articulated K-12 SEL competencies/standards; CASEL, n.d.-c), universal SEL programs are increasingly prevalent in school settings. In principle, these programs are intended to serve all students within the general education, Tier 1, settings; however, "universal" SEL programs are not without their faults.

Conceptually, SEL programs aim to take a public health approach to serving students. That is, students learn from a program and are able, in the short term, to develop SEL skills, and in the long term, apply these skills to contribute to a healthier functioning society (Greenberg et al., 2017). Indeed, most SEL programs are designed for universal implementation at the elementary level (Cipriano et al., 2023). However, much like within

the field of public health, inherent racism that is embedded into universal SEL programs can cause harm to minoritized individuals and communities (Bouffard, 2020; Mendez et al., 2021). Although universal SEL programs are implemented to reach all students, there is concern for whether students are learning from a culturally relevant curriculum (Mahfouz & Anothony-Stevens, 2020). Cultural relevance in education, when situated in a culturally relevant pedagogy (CRP) framework, has three goals: (1) promote student learning; (2) promote cultural humility; and (3) increase critical consciousness (Will, 2022). Cultural relevance in an academic curriculum has been shown to increase educational outcomes (Luevano & Collins, 2020). One area for improved cultural relevance is in SEL.

Universal SEL programs are often based on white middle-class norms (Hoffman, 2008). This leaves minoritized populations to fall behind their peers, face disproportionate punishment, or assimilate (i.e., adapt and lose their culture) (Ladson-Billings, 1995). Studies have shown that minoritized students who assimilate in higher-income settings have positive associations with academic achievement (Xie & Greenman, 2011), yet, this assimilation, or the accusation of assimilation by family members, can have harmful associations with anxiety and depression (Thornhill et al., 2021).

School personnel spend a great deal of time and resources each year implementing SEL programs with the hope that it yields beneficial outcomes. Yet, not all universal SEL programs are thoroughly investigated to understand their impacts on student academic and behavioral outcomes or to understand if the SEL program is culturally relevant for students. One example of this is the Be Good People program. The creators of this locally developed program have strived to make this program easily accessible, but the short and long-term impacts of the program are yet to be known.

Be Good People (BGP) SEL Program

BGP was developed by a group of educators during April of 2020. It was locally developed as a universal, modularized SEL program designed to serve students in grades K-12. The modules are developmentally written for the following grade bands: K-3, 4-5, 6-8, and 9-12. One significant benefit to BGP is that it is an open-access program (i.e., free to use). Indeed, nearly 2,000 schools/districts across the United States and the world have registered to use BGP (van Oss, 2024) demonstrating its reach.

The modularized style of program allows school districts the flexibility to create SEL schedules that support the needs of their students. Each lesson in the program is a complete module and is available as a Google Slide deck that takes 15-25 minutes to implement. The creators state that the modules can be taught in any order; however, they strongly recommend modules are organized in a sequence that is logical for the school and in related "clusters" of topics. An example kindergarten-level "cluster" could include the modules such as "Following Instructions," "Following Rules," "Using Active Listening," and "Getting the Teacher's Attention" that help transition students into school expectations for behaviors.

Although BGP does not have a formal evidence-base, the creators consider it to be at Every Student Succeeds Act's (ESSA's)-Tier 4 level of evidence. Tier 4 means the intervention "demonstrates a rationale" (Regional Educational Laboratory, 2019) via a well-defined logic model or theory of action based on rigorous research with an effort to study the effects of the intervention (California Department of Education, 2024; Regional Educational Laboratory, 2019). Importantly, each module of BGP is designed to address one of the five main CASEL SEL competencies of self-awareness, self-management, social awareness, relationship skills, and responsible decision-making (CASEL, n.d.-a).

Furthermore, BGP is aligned to a highly recommended SEL framework established by Durlak and colleagues (2011). Durlak et al. (2011) described how evidence-based SEL

programs fit the SAFE framework, which stands for sequenced (i.e., contain connected activities that follow skill development), active (i.e., using many learning and teaching strategies), focused (i.e., specified time to work on skills), and explicit (i.e., teaching specifically SEL skills). For example, a prominent SEL program in the literature aligned to the SAFE framework is SecondStep (Committee for Children, 2011). This program includes dedicated time for students to learn lessons at a developmentally appropriate level. BPG has a strong basis as the modules are structured to align with the SAFE framework that are essential for the effectiveness of a SEL program (Bond & Hauf, 2004).

Module Structure and Sequence. There are 68 to 92 primary modules depending on the grade band. Each module generally follows the structure of (1) stating module aims, (2) scenarios, and (3) discussion points. At the end of each module, students are shown what CASEL competency aligned skill they learned during the module. Additionally, there are extension activities that can be implemented on days when a module is not being delivered. The extension activities have a focal activity aligned with the module of the week. The community building activities include ice breakers, games, and mood meter temperature checks.

Each week, teachers have the option to implement extension activities throughout the week that are related to the week's primary module. Additionally, teachers also have the option to implement community building activities during the week. At the elementary level (i.e., K-3 and 4-5) modules are recommended to be delivered once per week, and potential extension and community building activities throughout the same week. However, at the secondary levels (i.e., 6-8, 9-12), example SEL sequences tend to rotate among the primary module, extension activities, and community building activities each week rather than implementing all three during a single week.

BGP is suitable to fit in a multi-tiered system of support structure where the first tier (i.e., Tier 1) is the universal tier that supports all students, the second tier (i.e., Tier 2) holds targeted supports, and the third tier (i.e., Tier 3) provides intensive support (Stoiber & Gettinger, 2015). The Core Edition of BGP is the universal level of the program and the Intervention Edition is built for Tier 2 and Tier 3 support. Given the reach of the Core Edition in the general education setting, the preventative focus of elementary level SEL programs, and the consistency of elementary module implementation, this study focuses on the Core Edition of BGP for elementary grade levels.

Current Study

The authors conducted an initial investigation of the BGP SEL program through a feasibility study, which is designed to determine whether an intervention is appropriate for further testing (Bowen et al., 2009). Feasibility studies are well-suited for novel interventions with no prior research. As such, the authors sought to better understand the feasibility of the program, with a particular focus on limited efficacy testing, practicality, and acceptability (Bowen et al., 2009). Limited efficacy testing examines if a new program demonstrates the potential to be successful, particularly for studies that leverage a convenience sample, shorter implementation periods, and limited statistical power (Bowen et al., 2009). Practicality investigates the extent interventions can be delivered with the available resources; acceptability focuses on whether the intervention is suitable for the implementers and those receiving the intervention (Bowen et al., 2009). That is, the authors sought to collect initial evidence of outcomes associated with the program, understand how it was being delivered in a naturalistic setting, and how users of BGP perceived the program. Despite the importance of understanding the relationship between locally developed, universal SEL programs that are being used to teach SEL skills to all students in numerous

districts, BGP has yet to have any of its implementation components or potential outcomes explored. The current study was guided by the following research questions:

1. To what extent do teachers implement BGP modules with fidelity (i.e., dosage and adherence)?
2. Are improved student social emotional skills observed over six months of implementation?
3. Are improved academic and behavioral student outcomes observed over six months of implementation (i.e., grades and social emotional behavioral screening data)?
4. What are teacher perceptions of the practicality of BGP?
    a. Practicality for this study was defined as the capability to implement the intervention (Bowen et al., 2009).
5. What are teacher perceptions of the acceptability of BGP?
    a. Acceptability for this study was defined as the suitability and satisfaction of the intervention (Bowen et al., 2009).
6. What are teacher perceptions of the cultural relevance of BGP?
7. What changes do teachers believe should be made to BGP (if any)?

A priori Hypotheses

Given the positive results on SEL skills and student outcomes found by Durlak et al. (2011) and Cipriano et al. (2023), the authors propose the following hypotheses to research questions two and three:

- If teachers implement Be Good People modules with high fidelity in terms of dosage and adherence, and students participate in the Be Good People modules with a high level of fidelity dosage (i.e., greater than or equal to 80%; Perepletchikova & Kazdin, 2005), then they will demonstrate positive growth in their social emotional skills observed from Time Point 1 (T1) to Time Point 2

(T2)

- If teachers implement Be Good People modules with high levels of fidelity in terms of dosage and adherence, and students participate in the Be Good People modules with a high level of fidelity dosage (i.e., greater than or equal to 80%), then they will demonstrate positive growth in their grades and an increase in social, emotional, behavioral screening data scores from Time Point 1 (T1) to Time Point 2 (T2).

Due to the exploratory nature of this study, no a priori hypotheses were posed for the other research questions. This study was pre-registered on OpenScience. Appropriate University IRB and school district approval was obtained prior to conducting the study.

Method

Research Context and Setting

This study was situated in a rural, public school district in the Midwestern United States. The school district's website characterized the values and mission of the district in the following manner: holistically supporting students in this school district includes developing and/or supporting students' academics, cognitive function, communication, daily living skills, English learning, physical needs (e.g., hearing, vision, sensory), motor skills, and social emotional learning. The mission includes enhancing student outcomes through data-based decision-making and evidence-based practices. District values include collaboration among educators and district partners, innovative and sustainable programs, and advocacy for educators and students at the individual and state legislature levels.

This research study was developed in partnership with district-level administrators, who were interested in understanding more about the implementation of BGP and student outcomes. The partnership was initiated in 2019 by district administrators, who sought consultation from Dr. Faith Miller in supporting SEL efforts in the district. Specifically, the

district administrators were interested in developing a Technical Assistance Center to build capacity for SEL implementation and sought consultation around best practices. Dr. Miller periodically consulted with the district administrators over a period of five years at the time the study was conducted, and facilitated the connection that would result in the current study.

Given the district's expansive and innovative work in SEL and Anna Li's research interest and experiences in SEL programs, Dr. Miller connected the district administrators with Anna to see if there were any district-based research needs regarding their SEL efforts. Further emails and meetings with the district SEL coordinator and a school principal allowed for the district and research team to communicate about their priorities in terms of research activities. The aims of the study were formed in conjunction with the SEL coordinator of the district. The SEL coordinator discussed wanting to understand: (1) what fidelity of BGP implementation looked like in the school, (2) if there was a relationship between the BGP program and improved student outcomes, and (3) understanding teacher perceptions of their training in BGP.

A single school was selected for study implementation in consultation with the district's SEL coordinator. For a school to be considered for inclusion, they needed to be implementing the Be Good People program in each grade level. Thus, schools and grades that primarily taught the Second Step (Committee for Children, 2011), the other main SEL program in the district, were not considered for participation. The SEL coordinator then asked school principals if they were interested in participating in the study. If the principal demonstrated interest, a formal meeting was held with the principal, the SEL coordinator, and members of our research team. During the formal meeting, the principal asked any clarifying questions they may have had.

Ultimately, an elementary school, serving approximately 400 students in grades K-6, was selected for study implementation. According to the school website, the school aims to develop a sense of confidence, knowledge, and responsibility to succeed in each of their students. There is also a positive behavior program with expectations of students to be, "respectful, responsible, and ready." These school-wide expectations are visible and reinforced through the school's classrooms, hallways, bathrooms, lunch time, recess, bus rides, and programs. In conjunction, these expectations are leveraged to support student learning and growth. This school was in its fourth year of implementation of the BGP program. Apart from fifth grade, all students from grades kindergarten to sixth grade implemented BGP as their only SEL program.

As of 2023, the federal designations for the enrolled student body by race and ethnicity was 81.5% white, 11.8% American Indian, 3.0% Hispanic or Latino, 1.5% two or more races, 1.3% Asian, 0.5% Black or African American, and 0.5% other Indigenous Peoples. In the school, 56% of the students accessed free/reduced-price meals, 15.5% received special education services, 3% were English language learners, and 0.5% were unhoused (Minnesota Report Card, 2023).

Participants

Recruitment

Participants were recruited using a convenience sampling method. Initial interest in study participation was first gauged by the school principal, who approached teachers who met the study qualifications (taught the BGP program) to determine interest in learning more about the study. After interest was gauged, the research team met with the interested teachers to discuss the details of the study, answer teacher questions, and obtain informed consent. University IRB (STUDY00020688) and district approval was obtained prior to data collection, and informed consent was provided by all participating teachers. Because non-

identifiable student data from the district was shared with the research team, the IRB determined that parental/caregiver consent was not necessary.

Ultimately, seven teachers signed up to participate in the study, including three first grade, one second grade, two third grade, and one fourth grade teacher. Descriptive demographic information was collected with a researcher-developed, teacher questionnaire. Teachers answered questions regarding their gender identity, race/ethnicity, levels of education, and years of teaching experience. Teachers were also asked if they supported the development of the BGP program, where support was defined as providing ideas, feedback, or joining meetings when creating the program. See Appendix E for the full questionnaire. The teachers who participated in this study had many years of teaching experience (M = 12, SD = 8), with most (n = 6, 86%) holding or pursuing a Master's degree. Two of the teachers (29%) supported the development of the BGP program. All the teachers identified their racial/ethnic identity as white (n = 11, 100%) and their gender identity as women (n = 11, 100%).

Student-level demographic data were collected via de-identified student demographics provided to researchers by the school administrators. There were a total of 111 students (N = 111) included in this study. The students were in first (n = 45, 40.5%), second (n = 14, 12.6%), third (n = 36, 32.4%), and fourth grade (n = 16, 14.4%) at the time of the study. Sex-assigned-at-birth ratios were relatively evenly split between female (n = 59, 53.2%) and male (n = 52, 46.5%). Racially/ethnically, the student sample was predominantly white (n = 90, 81.0%) with Indigenous students (n = 9, 8.1%) and multiracial students (n = 9, 8.1%) being the next largest racial/ethnic groups, and Hispanic students being the smallest group (n = 2, 1.8%). Only one student (0.9%) was identified as an English Language Learner, and over half the students were identified as receiving free or reduced

lunch (*n* = 64, 57.7%). Lastly, a small percentage of students were identified as having a disability (*n* = 16, 14.4%).

Measures

Weekly Teacher BGP Fidelity Log

Each week, teachers were tasked with completing a weekly fidelity log that measured: a) the dosage (i.e., exposure) of lesson modules and optional extension activities, along with b) the adherence to the lesson module (i.e., not extension activities) teachers taught. This log was developed on Google Forms. There were two dosage items and three adherence items. The dosage items asked, "How many lessons(s) did you implement?" and "How many extension activities did you implement this week?" Responses to the dosage items were provided in whole, numerical values. The adherence items were written in parallel to the observation form for instructional coaching that was made available on the BGP website. These items were 1) "To what extent were you able to lead through all of the slides?"; 2) "To what extent were you able to complete activities on all slides?"; 3) "To what extent did you have all students fill in learning logs at the end of the weekly lesson?". These items were rated on a 3-point scale where 0 = Not at all, 1 = Partial, and 2 = Fully. This log also tracked if there were any changes to the week's schedule such as not implementing the module at a typical time or drills interrupting modules. Teachers completed this log from the beginning of the study, referred to as T1 until the end of the study, referred to as T2. See Appendix F for the full measure.

Weekly Student BGP Learning Log

The purpose of this student learning log was to account for student level dosage (i.e., exposure) of the program. At the end of each lesson module, students crossed off either a "knowing yourself" (i.e., self-awareness) badge, "managing yourself" (i.e., self-management) badge, "thinking of others" (i.e., social awareness) badge, "getting along with others" (i.e.,

relationship skills) badge, or "making good choices" (i.e., responsible decision-making) badge that corresponded to the objective of the primary module in which they participated. This was a paper log where students also noted the date they earned the badge next to the badge. See Appendix G for an example page of the log.

Researcher BGP Observation Form

The researcher observation form was adapted from the BGP's observation form for instructional coaching (Be Good People, 2023). This form was completed for the fidelity checks of teacher adherence. Each item was rated on a 0 to 2 scale where 0 = Not Evident, 1 = Partially Evident, and 2 = Clearly Evident. See Appendix H for the full observation form questions.

Social Skills Improvement System - Social Emotional Learning Screener (SSIS–SEL)

The SSIS-SEL Screener (Elliot et al., 2018) is a teacher-reported measure where teachers rate their students' competencies in each of the CASEL-5 SEL competencies (i.e., self-awareness, self-management, social awareness, relationship skills, and responsible decision-making) on a 1 to 5, Likert-type scale. Descriptions of what can be observed of a student demonstrating skills at all levels are provided for teachers. In general, students with a "level 1" knowledge of the SEL skills demonstrate "very limited," "very rare," and/or "very poor" expressions of the skill. Students with a "level 5" knowledge of the SEL skills demonstrate "exceptional," "very strong," and/or "outstanding" expressions of the skill. In addition to the five CASEL subscales, this study also leveraged the Social Emotional (SE) composite score, which can range from five to 25, where students who score ten or less are considered "at-risk" socially, and students who score above ten are considered "not at-risk socially" (Elliot et al., 2018). With a correlation coefficient of .70 considered strong (Akoglu, 2018), the test-retest reliability coefficients for the self-awareness (r = .70), self-

management (*r* = .76), social awareness (*r* = .79), relationship skills (*r* = .80), and decision-making (*r* = .87) SE subscales were strong (Elliot et al., 2018).

In terms of validity evidence, the SSIS-SEL Screener holds some validity evidence. For correlational values with other similar measures such as the Social Skills Improvement System Performance Screening Guide (SSIS-PSG prosocial) (Gresham & Elliot, 2008), there were moderate correlational values for the self-awareness (*r* = .626), self-management (r = .634), social awareness (*r* = .626), relationship skills (*r* = .626) and decision-making skills (*r* = .757) subscales (Elliot et al., 2018). Additionally, when comparing the items of the SSIS-SEL SE composite scale to the CASEL-5 model, there was 100% agreement among five of six teachers for which SSIS-SEL items mapped onto which CASEL competencies (Elliot et al., 2018), demonstrating a high level of construct validity evidence.

Student Grades

The way students were graded at the school varied by grade level. Students in first grade were graded monthly in math and reading only. Both subject grades, each month, consisted of two to three standards (e.g., "I can spell new words by writing the sounds I hear," "I can tell if equations are true or false") each month. Standards under math and reading were rated by teachers on a 3-point scale where 1 = Not meeting standard (needs more practice), 2 = Making adequate progress (not quite there yet), and 3 = mastered the skill.

Second grade subjects that were included in analyses were handwriting, language arts, math, and spelling. They were also graded with various letter codes: E+ (Excellent +), E (Excellent), E - (Excellent -), S + (Satisfactory), S (Satisfactory), S - (Satisfactory -), or I (Improvement needed). The second-grade teacher stated that there is a level of personal interpretation in grading for when students are exceeding or meeting grade level standards,

but that she does assign letter grades based on a percentage chart. She considers "E's" as "A" equivalent, "S" as "B" and "C" equivalent, and "I" as "D" or below equivalent.

Third grade subjects that were included in the analyses were language arts, math, and music. Fourth grade analyses only included math topics (e.g., "Demonstrates fluency with multiplication facts") and music grades, as the other subject areas did not meet the criteria. Third and fourth graders were both graded on an A+ to F scale where A+ = 100%, A = 93-99%, A- = 92-90%, B+ = 87-89%, B = 83-86%, B- = 80-82%, C+ = 77-79%, C = 77-79%, C- = 70-72%, D+ = 67-69%, D = 63-66%, D- = 60-62%, and F = 59% or below.

my Social, Academic, and Emotional Behavior Risk Screener (mySAEBRS)

The mySAEBRS (von der Embse et al., 2016) measure is a 20-item, universal screener used to identify students who may be at risk for social, emotional, and/or behavioral problems. The measure is designed as a computer-based, self-report measure for students in grades 2-12. The broad scale (i.e., "total behavior") is composed of three subscales: social behavior, academic behavior, and emotional behavior. Analyses in this study were only conducted with the total behavior scale. A higher score on this measure indicates less concern for social, emotional, behavioral concern. The total score scale ranges from zero to 60 (FastBridge, 2023-a), with scores greater than or equal to 35 indicating "low-risk scores between 25 and 34 indicating, "some-risk", and scores between zero and 24 indicating "high-risk" (Fastbridge, 2023-b). The total behavior scale has demonstrated high internal consistency (Omega; $\omega = .828$) in a replication study by Kilgus et al. (2020), where acceptable internal consistency ω values are considered to be greater than .70 (McNeish, 2018). The total behavior scale has demonstrated adequate structural validity evidence for configural invariance, where each item loaded onto the same factors across each lower, middle, upper, and high-grade cluster (Kilgus et al., 2020). Kilgus et al. (2020) found the root mean square error of approximation (RMSEA = .08), standardized root mean

residual (SRMR = .07), the Comparative Fit Index (CFI = .85), and the Tucker-Lewis Index (TLI = .81) all indicated an acceptable to well-fitting model.

Social, Academic, and Emotional Behavior Risk Screener–Teacher Rating Scale (SAEBRS-TRS)

The SAEBRS-TRS (von der Embose et al., 2016) measure is a 19-item, universal screener used to identify students who may be at risk for social, emotional, and/or behavioral problems. This computer-based measure is available for teachers in grades K-12 to complete. The broad scale (i.e., "total behavior") is composed of three subscales: social behavior, academic behavior, and emotional behavior. Similarly to the mySAEBRS, a higher score on this measure indicates less concern for maladaptive social, emotional, behavioral problems. The total score ranges from zero to 57 (FastBridge, 2023-c), with scores greater than or equal to 37, indicating "low-risk, scores between 24 and 36 indicating "some-risk" and, scores between zero and 23 indicating "high-risk" (Fastbridge, 2023-b).

In a systematic review and quantitative synthesis, Goerdt (2023) found that the SAEBRS-TRS demonstrated promising technical adequacy in several areas. The weighted average internal consistency values for the total score scale ($\alpha = .92$) exceeded the acceptable level of $\alpha = .70$ (Cortina, 1993). Across 4 studies, Goerdt (2023) also reported strong, weighted test-retest reliability correlations for the total score scale ($r = .82$) where a correlation coefficient of .70 is considered strong (Dancy & Reidy, 2007). In terms of validity evidence, the SAEBRS-TRS has demonstrated both structural and concurrent validity evidence. The SAEBRS-TRS total behavior scale has demonstrated adequate construct validity with numerous other measures for social, behavioral, and academic risk. Goerdt (2023) reported a weighted average absolute correlation coefficient of .58.

Usage Rating Profile - Intervention, Revised (URP-IR)

Two facets of social validity were measured in this study: practicality and acceptability. The URP-IR (Chafouleas et al., 2011) was leveraged as a quantitative measure of these facets. This is a 29-item measure evaluating teachers' perceptions of six factors regarding the social validity of BGP. The six factors were (1) acceptability, (2) understanding, (3) home school collaboration, (4) feasibility, (5) system climate, and (6) system support. Each item is measured on a 6-point, Likert-type scale. The URP-IR was chosen as it demonstrates strong internal consistency reliability values for the six subscales (range = .72 to .95) (Briesch et al., 2013), where strong is considered a Cronbach's alpha (α) value above .70 (Netemeyer et al., 2003) The subscales of the URP-IR leveraged in this study were acceptability (α = .95) and feasibility (α = .84) (Briesch et al., 2013).

Culturally Responsive English Language Arts (CRE) Curriculum Scorecard

This CRE Curriculum Scorecard was designed by the New York University (NYU) metro center as a tool to aid educators and community members in determining how culturally responsive English Language Arts curriculum is that is used by school districts (Byran-Gooden et al., 2023). However, the authors of this scorecard have stated that it can be used for other subjects (Byren-Gooden et al., 2023, p. 12). This measure was included to quantitatively capture the cultural relevance of the BGP SEL program. Please see the "Key Definitions" section on "culturally responsive" versus "culturally relevant language." The developers of the scorecard recommend that the evaluation team is composed of three people with diverse identities and roles in the school and/or community.

This full scorecard encompasses several ratings. There are two quantitative tallies to be completed in the scorecard where raters go through the curriculum and tally the representation of characters and diversity of the authors. Raters tally for characters' race/ethnicity, gender, whether they have disabilities, are lesbian, gay, trans, queer,

intersex, asexual (LGBTQIA+), or if the characters are animals. There are also three scale scores that are provided on this measure: representation, social justice, and teachers materials. There is no total score for this measure. The representation score examines the extent of whom is reflected in the program. This score is composed of 17 items that refer to the character and author tally, the diversity of characters, and if those characters are portrayed accurately. The social justice score reflects the opportunities the program has for cultural responsiveness with eight items related to three categories: (1) decolonization, power, and privilege, (2) centering multiple perspectives, and (3) connecting learning to real life and action. Lastly, the ten items that comprise the teacher's materials score reflect the amount of guidance teachers have on culturally responsive teaching strategies.

All 35 items in the scorecard are rated on a +2 to -2 scale where +2 is very satisfied, +1 is satisfied, -1 is unclear, and -2 is not satisfied. Each scale's score is then categorized as either culturally destructive, culturally insufficient, emerging awareness, culturally aware, or culturally responsive. The full, 2023 CRE Curriculum Scorecard can be found by clicking on this link: https://steinhardt.nyu.edfu/sites/default/files/2023-05/CRE%20ELA%20Curriculum%20Scorecard%202023.pdf

Teacher Interview

The teacher interview was a semi-structured, 30-minute interview between the first author and each teacher individually. The full interview consisted of 19 questions split into three sections: (1) eight questions regarding general impressions, practicality and acceptability of the BGP program, (2) eight questions about the appropriateness in terms of the cultural relevance for students, and (3) three questions about teacher training in BGP. The researcher's interview script contained a shared definition for culture and Culturally Relevant Pedagogy. Culture was defined as, "Culture is passed down from one generation to the next. Things that can get passed down that encompass culture can include values,

norms, goals, items, language, and religion. Culture can affect our behavior and our perceptions of others' behaviors," (Fiske, 2002; Spencer-Oatey, 2012) and Culturally Relevant Pedagogy was defined as, "a model that supports student academics, learning of their own and others' cultures, developing an understanding of social, political, and economic contradictions and how to take action against oppression" (Ladson-Billings-1995).

Procedure

Teacher Training

At the research site's school, teachers were provided informal training where the district's SEL coordinator showed the teachers the BGP website, resources, and how to access different materials and then answered questions. There were also training videos on the BGP website that are made freely available for teachers in each grade band. The videos are hosted by the creator and primary author of BGP Nic van Oss as he demonstrates an example module. Teachers at the school were not required to watch these videos.

Data Collection

This study took place over the course of the 2023-2024 academic school year. Teacher participants and our research team met at an initial meeting in December 2023, just prior to the district's winter break. At this meeting, teachers were instructed on how to access and complete several measures. Measures that were completed during the first meeting were the TDQ, SSIS-SEL, SAEBRS-TRS, and the URP-IR. mySAEBRS was already implemented as standard practice at this school for the second, third, and fourth grade teachers. Thus, it was not necessary for teachers to be instructed by the research team on how and when to complete the mySAEBRS assessments. The teachers were then instructed to fill out the weekly teacher fidelity log at the end of each week until the end of the study. Data collection of the weekly fidelity data began halfway through the second quarter of the

academic year (i.e., after winter break) and ended near the end of May 2024, approximately two weeks before the end of the academic year. The second set of measures were sent via email to the teachers in a document that provided links to the SSIS-SEL and URP-IR. Teachers were then also reminded to complete the SAEBRS-TRS. Teachers were then all interviewed individually for 30 minutes in the second half of May 2024. All interviews were conducted virtually on Google Meet. Interviews were audio recorded and transcribed using Otter.AI software (Otter.ai, 2024).

BGP Program and Research Implementation

Teachers taught one primary module of the BGP program to their students once a week, as usual. At the end of each module, teachers held time for students to complete their learning log and filled out the weekly teacher fidelity log. Researchers conducted at least two full observations of module implementation for each teacher during the academic year. Teacher measures were completed at two time points, once during the second quarter of the academic year and once near the end of the academic year. Each measure, except for the teacher demographic questionnaire, teacher/student fidelity logs, and interview, was completed twice. From December 2023 to January 2024, teachers completed all measures for the first time, excluding the weekly fidelity log and interview. In May 2024, teachers completed all measures, excluding the demographic questionnaire and weekly fidelity log. Teachers were compensated with a $50 gift card for completing the first set of T1 measures, and then compensated with another $50 gift card for completing the T2 measures.

Planned Data Analyses

All statistical analyses were conducted in R version 4.2.2 (R Core Team, 2022) and all interviews were transcribed.

Research Question 1: Teacher Fidelity Levels

Descriptive analyses were performed to understand two facets of fidelity: dosage (i.e., the number of modules taught) and adherence. Data from the weekly teacher fidelity log were leveraged for the dosage and adherence of primary BGP module implementation (i.e., not the extension activities or the community building activities). Dosage was calculated by dividing the total number of implemented modules by the total number of scheduled modules to obtain a final dosage percentage for each teacher. A weighted average of the dosage percentages was then calculated. An adherence percentage was also calculated for teachers leading through all the slides and completing all the activities per module. Both percentages were calculated by dividing the total points scored by total points possible. A weighted average of the adherence percentage for both the slides and activities completion were calculated. Fidelity checks for teacher adherence to completing all module slides and activities were also conducted by the two researchers. Three of these fidelity checks were conducted per teacher for a total of 21 researcher fidelity checks. Interrater reliability (IRR) was also calculated between the two researchers' observations on two separate occasions.

Student-level dosage was calculated via the weekly student learning logs. The number of modules students were present for was divided by the number of modules where teachers had students record their presence to gain a dosage percentage. An average percentage of modules for which students were present was calculated for each classroom. A weighted average was then calculated for overall student-level dosage.

Research Question 2: SEL Skills Learned (SSIS-SEL)

Given the small sample size and potential violations to parametric assumptions, a paired, Wilcoxon signed rank test was chosen as the analysis method for each subscale of the SSIS-SEL (i.e., each of the five SEL domains). Assumptions were then checked for each

subscale. Students' scores in each subscale were compared from T1 to T2. A paired Wilcoxon signed rank test effect size, a correlation coefficient, was also calculated to contextualize the results (Tomczak & Tomczak-Łukaszewska, 2014). The Wilcoxon effect sizes were interpreted using the guidelines articulated by Cohen (1992) such that .10 is a small effect, .30 is a moderate effect, and .50 is a large effect (Cohen, 1992; Tomczak & Tomczak-Łukaszewska, 2014).

Research Question 3: Student Academic and Behavioral Outcomes

Given the small sample size and potential violations to parametric assumptions, a paired Wilcoxon signed rank test was chosen as the analysis method for analyzing student grades and behavioral outcomes through the following measures: (1) student grades, (2) mySAEBRS, and (3) SAEBRS-TRS Wilcoxon signed rank test assumptions were checked for all data prior to running analyses.

Academic Outcomes. First, the second-grade student grades were re-coded on an incremental scale of one, from lowest score to highest score, where 1 was the lowest score and 8 was the highest score. The third and fourth grade grades were re-coded similarly where 1 was the lowest score and 13 was the highest score. See Table 11 for the full, re-coded scale.

Given the various grading methods, academic outcomes were analyzed leveraging separate, paired Wilcoxon signed rank tests for each grade level. The T1 academic grade scores for each first-grade student were composed of the average of their February reading and math grades. The T2 academic grade scores for each first-grade student were composed of their averaged May reading and math grades.

Second, third, and fourth grade student academic performance was measured by their second (Q2) and fourth (Q4) quarter report card scores that corresponded to T1 and T2 points of this study, respectively. There were approximately 40 days in each academic

quarter. Subjects that were identified by the researchers as "graded" met the following criteria: a) the subjects had a Q2 and Q4 grade and b) students received unique scores. For example, subjects where students received a pass/fail grade or all students automatically received the highest grade (e.g., physical education) were excluded. A total score was created across academic subjects for all grades because SEL has been shown, in practice, to impact how well students do in all subject areas (Cipriano et al., 2023). A paired Wilcoxon signed rank test effect size was also calculated to contextualize the results at each grade level.

Behavioral Outcomes. For the mySAEBRS and SAEBRS data, a paired Wilcoxon signed rank test was again administered as the data did not follow a normal distribution. The mySAEBRS and SAEBRS total behavior scale scores were compared for each student from the school's late fall screening period to the spring screening period, which corresponded to T1 and T2, respectively. Subscale analyses were not run, as the total behavior scale has been reported with the strongest internal consistency (Kilgus et al., 2020).

Research Question 4: Teacher Practicality Perceptions of BGP

The authors leveraged two data sources to better understand teachers' practicality perceptions of BGP: the feasibility subscale of the URP-IR and the teacher interview. The items in the feasibility subscale were summed. For the teacher interview, the qualitative, thematic analysis approach was leveraged to answer this question and the following research questions. Thematic analysis is a strategy described by Braun and Clarke (2012). The six steps described by Braun and Clarke (2012) were taken to analyze the transcripts from the interviews: 1) familiarizing yourself with the data, 2) generating initial codes, 3) searching for themes, 4) reviewing potential themes, 5) defining and naming the themes, and 6) producing the report. An inductive approach was taken to analyze this data by

creating themes from the data presented. The interview questions thematically analyzed for responses that best fit this study's definition of practicality were from the first set of interview questions (Questions 1-2 and 5-8; see Appendix I).

Research Question 5: Teacher Acceptability Perceptions of BGP

The authors again leveraged two data sources to better understand teachers' acceptability perceptions of BGP: the acceptability subscale of the URP-IR and the teacher interview. The items in the acceptability subscale were summed. Interview questions 1-2 and 5-8 were again thematically analyzed for responses that best fit with this study's definition of acceptability.

Research Question 6: Cultural Relevance Perceptions of BGP

Quantitative Analysis. The CRE Curriculum Scorecard was leveraged for the quantitative analysis of this question. From each grade level K-5, five, unique modules representing each SEL competency were selected for analysis. A total of 30 modules were analyzed.

The scorecard was completed by the first author of this study, a Ph.D. candidate in school psychology and an additional Ph.D. student who also conducted the researcher observations IRR, and the SEL coordinator. All three individuals familiarized themselves with the questions and scoring of the tool. The two Ph.D. students completed the representation of characters' tally together. Although the scorecard states tallying characters from all images in the curriculum, only characters from the scenarios were counted, and the images in the scenario were not used in the tally for character representation. The researchers decided to not take race, ethnicity, gender, disability, or sexual orientation into account via the images for two reasons: (1) the images were general stock photos that may not have exactly matched the scenario (e.g., two friends mentioned, but only one friend pictured), and (2) assuming individuals identities from an image allows

for a wide range of possibly incorrect interpretations and discounts identities that may not be visible. Thus, only the gender pronouns in the written scenario text were counted. See Appendix J for an example of how the tally was conducted. The two Ph.D. students then independently completed the Representation score section of the report card. After completing their individual ratings, they built consensus on what each item should be scored.

The first author and the SEL coordinator for the school district then conducted the author representation tally with knowledge from an individual who was part of the BGP development team. The same process that was taken for the representation score was then followed for the Social Justice and Teacher Materials section of the scorecard. The teacher materials that were referenced for scoring were the teacher training videos for the K-3 Band Core Edition (St. Croix River Education District, 2022b), the 4-5 band Core Edition (St. Croix River Education District, 2022c), and the video titled "Creating a School-Wide Schedule for SEL Instruction (Elementary Schools)" (St. Croix River Education District, 2022a). These training videos were chosen as they were located under the "Make it Happen" tabs of the BGP webpage.

Qualitative Analysis. The teacher interview was leveraged for further understanding of the cultural relevance of BGP. Questions from the teacher interview that could provide the cultural context of the study area and the teachers along with their perceptions of the cultural relevance of the BGP program were identified. These questions came from the second section of the teacher interview (Questions 9-13, and 15, see Appendix I) and were thematically analyzed. The questions included: (9) "Given this definition of culture, with a particular focus on the aspects that most impact your daily life, how would you describe your culture?"; (10) "How would you say the description of your culture is similar to or differs from the culture of your students?"; (11) "How would you say

your culture is reflected in the Be Good People program, if at all? (12) "How would you say all of your students' cultures are reflected in the Be Good People program, if at all?"; (13) "In what ways is the Be Good People program made for individuals like your students, if at all?"; and?" (15) With this shared understanding of Culturally Relevant Pedagogy, have you received any training in Culturally Relevant Pedagogy in SEL or otherwise? If so, what did it look like?

Research Question 7: Changes Teachers Would Make (If Any)

To better understand what changes teachers would make to BGP, questions from each section regarding changes were thematically analyzed. The questions analyzed were: (3) "How would you change the Be Good People program to better suit your needs"; (4)"How would you change the Be Good People program to better suit your students' needs, if at all?'; (14) "In what ways would you change the program to better suit your students' cultures, if at all?"; (16) "How do you make changes to the BGP curriculum to align with culturally relevant pedagogy, if at all?"; and (19) "How would you change the Be Good People training, if at all?"

Missing Data

If less than 10% of data was missing in any data set, listwise deletion was used (Bennett, 2001). If more than 10% of data was missing for any measure, incomplete and/or fully missing data were removed and appropriate caution in interpreting the results were made. In sum, teachers provided weekly teacher fidelity log data for 127 of the 147 (86%) modules taught. Additionally, most classrooms returned all of their student learning logs, with 100 of 111 student logs returned (90.1%). However, student level dosage varied, with only two classrooms providing student dosage data for 100% of the modules they taught. The other five classrooms had missing student dosage data missing for 10% to 35% of the modules. Thus, calculations of student-level dosage were only made based on weeks where

teachers had students fill out the student learning logs and should be interpreted with caution.

Due to missing data from either a T1 or T2 rating, researchers removed one participant (1%) from the SSIS-SEL self-awareness subscale analysis, two participants (2%) from the relationship skills analysis, and one participant (1%) from the responsible decision-making subscale analysis. For academic grades, one of the 45 first grade students (2.2%) was removed from the analysis as they left the school district in April 2024 and did not have any grades for May 2024. The fourth-grade analysis had six of fifteen students removed (40%) due to non-reported grades.

Only second, third, and fourth graders were eligible to complete the mySAEBRS. Across these three grades, 54 of the 66 (82%) students completed the mySAEBRS screener in both fall and spring. Thus, 12 of the 68 (12%) students' data were missing. One teacher did not complete both T1 and T2 SAEBRS-TRS ratings. Thus, 20 students were missing from the sample (18%). The final analytic sample contained 91 of the 111 students (82%).

Exploratory Data Analyses

The SSIS-SEL, SAEBRS-TRS, and mySAEBRS provide indicators for maladaptive behavioral risk level. The SSIS-SEL SE Composite is split into two categories: "no-risk" and "at risk.", while the SAEBRS-TRS and mySAEBRS total scores provide "low", "some" and "high" risk categories. Given the categorical nature of this data, to understand differences in distribution among the categories from T1 to T2, a chi-square analysis was conducted for the risk status based on the total score of all three measures. Results of each chi-square are provided after the planned data analyses.

Results

Research Question 1: Teacher Fidelity Levels

Dosage

Based on the teachers' self-reports, all but one teacher implemented over 80% of the expected 21 modules (*M* = 86%, *SD* = .10; Susan = 100%, Julia = 81%, Ava = 71%, Laura = 95%, Christina = 81%, Rebecca = 90%, Natalie = 86%). The weighted average for the dosage of the modules was 87%. Please note: all names are pseudonyms created by the authors.

Adherence

The teachers' self-reports indicated they completed almost all of the slides during each module (*M* = 98%, *SD* = .04) and almost all the activities of each module (*M* = 90%, *SD* = .09). The weighted averages of the completed slides (98%) and completed activities (91%) were the same or nearly the same as the non-weighted average percentages.

Researcher Fidelity Checks

Although researchers intended to complete IRR observations on 21 modules, only 86% (*n* = 18) of the 21 scheduled observations were complete. Of these observations, teachers and researchers had 100% agreement on the BGP slides adherence and 61% agreement on the activities adherence.

IRR

Of the two observations where the researchers conducted IRR agreements with each other, there was 100% agreement on the BGP slides adherence and 100% agreement on the activities adherence.

Student Learning Logs

Knowing the true number of lessons students were present for was difficult as greater than 10% of data was missing for five of the seven classrooms. Of the weekly

modules that contained student-level dosage, on average, students were present for greater than 80% of the modules taught (*M* = 88%, *SD* = .05; Susan's Classroom = 88%, Julia's Classroom = 89%, Ava's Classroom = 88%, Laura's Classroom = 82%, Christina's Classroom = 90%, Rebecca's Classroom = 97%, Natalie's Classroom = 82%). The weighted, average percentage for student level dosage was 88%. However, caution should be taken when considering if there was a high-student dose level given the large number of modules (between 10-35%) missing student-dosage information.

Research Question 2: SEL Skills Learned (SSIS-SEL)

Hypothesis 1 - SEL Skills Will Increase in a Positive Direction with Greater than 80% Teacher and Student Level Fidelity

With greater than 80% teacher fidelity and student fidelity, on average, this hypothesis was partially supported. Of the five SSIS-SEL subscales, three returned statistically significant differences in the desired direction from T1 to T2: self-awareness, self-management, and relationship skills. Teacher ratings of students' self-awareness skills increased from T1 (*MDN* = 3, *IQR* = 3, 4; *M* = 3.23, *SD* = .97) to T2 (*MDN* = 4, *IQR* = 3, 4; *M* = 3.61, *SD* = .94) of the study. The change in teacher ratings for students' self-awareness skills represented a statistically significant change with a moderate effect size (W = 448, z = -3.81, $p < .0001$, r = .355). The change in teachers' ratings of students' self-management skills also increased statistically significantly with a small effect (W = 372, z = -2.78, p = .006, r = .27) from T1 (*MDN* = 3, *IQR* = 2, 4; *M* = 3.17, *SD* = 1.06) to T2 (*MDN* = 4, *IQR* = 2, 4; *M* = 3.40, *SD* = 1.12). Lastly, teachers' ratings of students' relationship skills increased statistically significantly with a small effect size (W = 294, z = -2.09, p = .04, r = .21) from T1 (*MDN* = 3, *IQR* = 3, 4; *M* = 3.48, *SD* = .99) to T2 (*MDN* = 4, *IQR* = 3, 4; *M* = 3.64, *SD* = 1).

Two of the SSIS-SEL subscales showed no statistically significant differences between T1 and T2: social awareness and responsible decision-making. Teachers' ratings of

students' social awareness skills did not statistically significantly change (W = 708, z = -.32, p = .75, r = .05) between T1 (*MDN* = 4, *IQR* = 3, 4; *M* = 3.69, *SD* = .94) and T2 (*MDN* = 4, *IQR* = 3, 4; *M* = 3.72, *SD* = .98). Ratings of students' responsible decision-making skills (W = 801, z = -0.94, p = .35, r = .10) also did not change statistically significantly between T1 (*MDN* = 4, *IQR* = 3, 4; *M* = 3.48, *SD* = 1) to T2 (*MDN* = 4, *IQR* = 3, 4; *M* = 3.56, *SD* = 1.07).

Descriptively, a large majority of students fell in the no-risk category based on the SE Composite at both T1 (n = 101, 93.5%) and T2 (n = 95.4%) as opposed to the at-risk category at T1 (n = 7, 6.5%) and T2 (n = 5, 4.6%). The exploratory chi square analysis demonstrated there was not a statistically significant difference in the number of no-risk and at-risk students between T1 and T2 (X^2 (1, N = 91) = .09, p = .77).

Research Question 3: Student Academic and Behavioral Outcomes

Hypothesis 2 - Student Academic and Behavioral Outcomes Will Improve with Greater than 80% Teacher and Student Level Fidelity

With greater than 80% teacher and student fidelity on average, this hypothesis was partially supported. Student academic performance in terms of grades increased statistically significantly for most grade levels. First grade students' academic performance increased statistically significantly with a moderate effect size (W = 164, z = -2.29, p = .02, r = .38) from T1 (*MDN* = 2.71, *IQR* = 2.32, 2.86; *M* = 2.54, *SD* = .52) to T2 (*MDN* = 3, *IQR* = 2.5, 3; *M* = 2.66, *SD* = .54). Second grade students' academic performance increased statistically significantly with a large effect size (W = 8.5, z = -2.19, p = .03, r = .54) from T1 (*MDN* = 18, *IQR* = 12.75, 22.7; *M* = 18, *SD* = 5.75) to T2 (*MDN* = 19, *IQR* = 16.25, 23; *M* = 19.79, *SD* = 5.98). Third grade students' academic performance also increased statistically significantly with a moderate effect size (W = 128, z = -2.38, p = .02, r = .42) from T1 (*MDN* = 32, *IQR* = 28, 35; *M* = 30.72, *SD* = 5.73) to T2 (*MDN* = 32.50, *IQR* = 29.50, 35; *M* = 31.75, *SD* = 5.21). Academic performance for fourth grade students did not change statistically significantly

(W = 20, z = -0.31, p = 0.75, r =.10) from T1(*MDN* = 49, *IQR* = 49, 52; *M* = 50.22, *SD* = 1.72) to T2. (*MDN* = 51, *IQR* = 50, 51; *M* = 50.33, *SD* = 1.32)

The second hypothesis was not supported for social, emotional, and behavioral screening data. Students' self-reported total scores on the mySAEBRS screener did not increase statistically significantly (W = 656, z = 0.17, p = 0.86, r = .02) from T1 (*MDN* = 40.50, *IQR* = 33.25, 47;75; *M* = 39.83, *SD* = 9.73) to T2 (*MDN* = 40.50, *IQR* = 33, 44; *M* = 39.11, *SD* =7.85). Teachers' ratings on the SAEBRS-TRS total score also did not increase statistically significantly (W = 1454, z = -.05, p = .97, r =.003) from T1 (*MDN* = 45, *IQR* = 36.50, 52; *M* = 43.41, *SD* =10.33) to T2 (*MDN* = 46, *IQR* = 37, 52; *M* = 43.34, *SD* =10.93).

Descriptively, most students were in the low-risk category at T1 and T2 for both the SAEBRS-TRS (T1 low-risk: n = 68, 74.7%; T2 low-risk: n = 69, 75.8%) and mySAEBRS (T1 low-risk: n = 37, 68.5%; T2 low-risk: n = 36, 66.7%) measures based on the total scores. Additionally, there were similar rates of students at the some-risk and high-risk categories for both the SAEBRS-TRS (T1 some-risk: n = 19, 20.9%; T2 some-risk: n = 17, 18.7%; T1 high-risk: n = 4, 4.4%; T2 high-risk: n = 5, 5.5%) and mySAEBRS (T1 some-risk: n = 13, 24.1%; T2 some-risk: n = 16, 29.6%; T1 high-risk: n = 4, 7.4%; T2 high-risk: n = 2, 3.7%) measures. The exploratory chi-square analysis for the mySAEBRS (X^2 (2, N = 54) = .99, p = .61) and SAEBRS-TRS (X^2 (2, N = 91) = .23, p = .89) did not demonstrate a statistically significant difference among the students in the low-risk, some-risk, and high-risk categories.

Research Question 4: Teacher Practicality Perceptions of BGP

Teachers' perceptions of the feasibility of BGP were high (*M* = 30.7, *SD* = 3.59). The highest possible summed rating a teacher could provide on the subscale was a 36. Teachers' ratings ranged from 25 to 35. The lowest rated item (a reverse-scored item) stated, "This item is too complex to carry out accurately" (*M* = 3.71, *SD* = 1.60). All other items had an

average rating above 5 which indicated most teachers "agreed" to "strongly agreed" with the statements. The highest rated item, on average, was item number 13 that stated, "Preparation of materials needed for this intervention would be minimal (M = 5.86, SD = .38). Then, using information obtained from the teacher interview, one central theme regarding the ease of access to the program was identified for teacher practicality perceptions.

Theme 1: BGP is Convenient for Teachers

First and foremost, teachers enjoyed having a pre-made program that was accessible online. Others shared how the SEL coordinator prepared a whole-year scope and sequence of modules to teach each week of the academic year that allowed teachers to know which modules to teach. One teacher stated, "everything is prepared, ready to go…, it's linked in, and we just click, you know, click on our lesson plans and it's there" (Susan). Another teacher echoed, "scope and sequence is set up for me, and I know that if I teach it I'm covering all the SEL standards" (Ava). One teacher addressed how BGP overcome barriers that other schools see for implementing SEL programs.

> "It's like a 'done for you' curriculum. You know, with slideshows that we can link in like it's all digital. You know, we've done other curriculums where you need a book or outside of materials, and it's easy to like, misplace those. It's easy to not have them ready for a sub" (Julia*).

Table 12 provides a joint display of the quantitative and qualitative data germane to this research question. The joint display triangulates information to have a stronger understanding of teachers' practicality perceptions of BGP. Although teacher ratings generally ranged from agreed to strongly agreed on each URP-IR Feasibility subscale item, there were both quotes that both supported and contradicted teachers' ratings.

Research Question 5: Teacher Acceptability Perceptions of BGP

Teachers' perceptions of the acceptability of BGP based on the URP-IR rating scales was fairly high when summing their scores (M = 40.57, SD = 7.09) given the maximum score teachers could rate the subscale was a 54. However, there was a larger range (low = 31, high = 48) of ratings compared to the feasibility subscale, demonstrating more varying opinions on BGP's acceptability level compared to its feasibility level. Average ratings across teachers for the acceptability subscale (M = 4.5, SD = .79) indicated that teachers tended to be between "slightly agree" and "agree" for each item they rated. The lowest-rated item (a reverse-scored item) stated, "I would not be interested in implementing this intervention (M = 3.86, SD = 1.07). The highest-rated item stated, "I would implement this intervention with a good deal of enthusiasm" (M = 5.14, SD = 1.07).

There were also two themes extracted from the teacher interviews regarding the acceptability of BGP to the teachers and students. The themes centered around finding BGP to be applicable to the students' lives, yet the level of applicability was dependent by grade level.

Theme 2: BGP is Pertinent for Students

In general, teachers noted that BGP was relevant and important to students' lives. Several teachers identified that the skills discussed in BGP are present in the students' daily lives. One teacher stated, "every lesson that comes up, we've seen…in the classroom or they've experienced" (Laura). Another teacher stated,

> "I think [BGP is] very relevant because a lot of the issues or skills that we're talking about they face, and their lives that they're dealing with for social and as well as their emotional part of their lives, such as with sportsmanship. They're now breaking away from adults always leading their lives in third grade to more of that maturity to peers. And they're really in that conflict. So, like, even sportsmanship is

> a big one, skill that we teach. And it's very relevant. None of them come in with very good sportsmanship at all. They face a lot of issues with that and Be Good People really address it" (Rebecca*).

When it came to the importance of BGP, teachers resoundingly agreed that the skills taught in BGP were important. One stated that, "it really does help them navigate their emotions and their feelings and just experiences that you go through everyday life at their age range" (Rebecca*). Another found that BGP was something helpful to teach for the increasing students' social, emotional, and behavioral skill needs.

> "I mean SEL is a hugely, hugely important thing happening. I, you know, I see kids struggle more and more with that, but that may not necessarily be just because teachers are not doing something, or home may not be doing something. You know, I think there's a whole bunch of reasons why kids seem to be struggling more with social emotional skills, but definitely very, very relevant to their lives. Like we ought to be teaching something I think to fill the gap" (Julia*).

Although many of the teachers found BGP to be meaningful to student's lives, teachers also found that the level of concepts may have been developmentally inappropriate for students.

Theme 3: BGP Skill Generalizability is Limited at the Elementary Level

According to teachers interviewed, students' abilities to generalize the BGP skills from the lesson to real life situations seemed to be limited for various reasons at the elementary level. Only one teacher (third grade) reported that they see students using the BGP skills at school and have heard from families that their students have improved SEL skills at home. The teacher said in places with unstructured times like recess, she can,

> "Really see the maturity especially in this grade level, at least for me, with just expressing their emotions and communicating their needs and their wants better. Um, yeah, you just see a lot of emotional growth throughout the year which I think is

> great and just them taking their time and using those kind words, and they're listening before speaking and acting, which is really cool to see" (Rebecca*).

However, this teacher's experience appears to differ from other teachers' observations. First grade teachers suggested that the program might not be suitable for their students' developmental levels. Beyond feeling they need to simplify some of the language in the modules, they reported rarely seeing students using BGP skills, and if they did, the students needed extra prompting from the teacher to use the skills. For example, at the first-grade level, a teacher noted that the skills discussed were,

> "Beyond their age level at this point. Okay. Some of those things, I think they're more… 'me' centered. So, it's harder for them to look beyond themselves. And so some of those [scenarios] that are... where they have to kind of put themselves in another person's...from another perspective, it's harder for them because they're still sort of centered around their world and what they're doing" (Susan*).

Similarly, teachers at the older grade levels discussed difficulty applying the skills learned in moments where students could most benefit from in the moment and prompting is necessary to get students to use their skills. A third-grade teacher shared,

> "I don't know how much they themselves refer to it. But I will say things at recess like, 'oh, make sure you're using good sportsmanship,' or we were coming in from recess and someone had cheated or something, and I said, 'oh, was that good sportsmanship?' And they'll say, 'no, we just talked about this morning, so I can't believe they're not doing it.' So, it's like [they] remember it, but I don't know how much they're… I don't know… they know we talked about it, but it doesn't necessarily translate to actual change" (Natalie*).

Despite the need for simplification and prompting and limited application at the students' current grade level, teachers agreed that BGP lays a foundation for future social, emotional,

and behavioral development. The second-grade teacher in the study recognized the limited applicability of the BGP skills for elementary students but noted,

> "Not all of [the skills] may be able to apply the skills to real life yet. They are, I think, building their toolbox and hopefully when they are older, high school, college adulthood, they have appropriate ways to interact with and respond to others, and that they're aware of themselves and others. So, I think like a lot of us growing up, we weren't taught these skills and how our emotions and how things affect us. And so maybe they're not able to apply it yet but hopefully, you know, that every year we're drilling it into their head, you know, like, 'Hey, be aware, look at this,' and that...hopefully they'll be able to apply it to their lives in the future" (Laura).

Table 13 provides a joint display of the quantitative and qualitative data germane to this research question. The joint display triangulates information to have a stronger understanding of teachers' acceptability perceptions of BGP. Teacher ratings generally ranged from slightly agree to agree on each URP-IR Acceptability subscale item. Although there were few quotes that related to both the themes identified and the items of the URP-IR Acceptability subscale, there were both supporting and contradicting quotes found in relation to each of the items.

Research Question 6: Cultural Relevance Perceptions of BGP

CRE Curriculum Scorecard Results

In total, 303 characters were mentioned in the 30 modules that were rated. 302 (99.7%) of the characters were human and one of the characters was an animal (0.3%). All the human characters were racially ambiguous with approximately half of the human characters using gender pronouns that identified them as girls/women (*n* = 95, 31%) or boys/men (*n* = 88, 29%). The other portion of human characters had unknown gender identities (*n* = 119, 40%). BGP had five individuals who primarily developed the curriculum

who represented white women (*n* = 4, 80%) and one white man (*n* = 1, 20%). The representation score was a -8 which falls in the "culturally insufficient" range, the social justice score was a -6, which also falls in the "culturally insufficient" range, and the teacher materials score was a -4, which also fell in the "culturally insufficient" range.

Culturally insufficient was defined in different ways across the three scoring categories. A culturally insufficient representation score means, "The curriculum likely has culturally and racially ambiguous characters. Few characters and stories are portrayed in a culturally and historically accurate way. There is likely little to no diversity among curriculum contributors and illustrators (Byran-Gooden et al., 2023, p. 23). For the social justice score, culturally insufficient describes how, "The curriculum predominantly centers white or Eurocentric ideas and culture in most of its components. For the most part students are not encouraged to think critically or take action to combat inequity. The curriculum provides weak connections to students' lived experiences. There is hardly any opportunity for teachers to engage cultural responsiveness" (Byran-Gooden et al., 2023, p. 24). Lastly, a culturally insufficient rating for the teachers' materials deems "There is a little guidance on engaging diverse learners or culturally responsive teaching in the teacher's materials, but it is mostly on a superficial or symbolic level. It is seen as additive, rather than central to the curriculum and teaching" (Byran-Gooden et al., 2023, p. 25). See Table 14 for exemplary quotes from teachers that supported the results of the CRE Curriculum scorecard scoring.

Cultural Context of the Teachers and Study Locale

Prior to sharing the qualitative thematic results directly related to teacher perceptions of the cultural relevance of the BGP program, it is important to understand the cultural context in which this study took place. The authors first provide context surrounding teachers' demographics, the study locale, teachers' training in CRP, and

teachers' cultural backgrounds that might have shaped teachers' cultural relevance perceptions of BGP. *Please note that all the following perspectives are only from the opinions of the teachers interviewed in the study, and they do not necessarily represent the opinions or views of all teachers, the school, the district, or the local area. Additionally, all names have been replaced with pseudonyms. If a name is followed by an asterisk (*), that indicates the direct quote has been edited for clarity by removing filler words, linguistic utterances, and/or repeated words; however, the integrity of the quote was not changed.*

Teacher Demographics and Study Locale. First, the demographics reported by the teachers and some of their interview responses when describing their cultural backgrounds indicated that perspectives shared in this study came from a relatively homogenous group. In the interview, three teachers mentioned their race (i.e., white), religion (i.e., Christian), and language (i.e., English) as part of their culture. Moreover, a teacher recognized that her students shared many of her identities as they were raised in similar contexts. She noted,

> "I also grew up in a small town here in [state name]. Majority of the people I went to school with were white. So, we celebrate the same holidays. We have the same religions… we have a lot of the same, like, you know, there's poverty and drug use and stuff in the homes…, I went to school with kids that also had the same issues that these kids have going on in their homes. And that doesn't make them bad kids or nothing, you know what I mean? So, just knowing that kind of what they're going through, and what they're seeing on a daily basis, and what they view as normal, you know, can help me to understand their, their situation or where they're coming from or how they think" (Laura).

Further, one teacher shared her perspective on the local history and area after growing up where this study was conducted. When asked if she made changes related to culturally

relevant pedagogy, she provided insight into local tensions. The teacher described the importance for her to teach meaningfully to her students.

> "I really tend to just teach from my, my core values of what I feel is right. And so, I think…it's important... what I didn't learn about, like, any of the real history of our country until college, and it was shocking to me, you know, and then I did…I told you, I studied Latin American history and got really into that. And similar types of things happened there. Right. So, with the schools and the disappearing and all that stuff was so heavy for me, and I just remember having no idea until I was, a sophomore in college. No idea. So, I, really, that's important to me" (Julia*).

She then described tensions between a nearby school district and the Indigenous population as the district voted against allowing the Indigenous community to perform a drum song at the high school graduation ceremony. The teacher then delved into the intersections of economic class and race,

> "I'm jumping subjects, but our Native American students are coming from a really low socio-economic status, right. But most of our white kids are too. So, it's really hard for our white population to have this...this empathy for the Native American students when they're living in the same situation. Like, both are living in just the shit of the shit… and they have…both have just generational trauma that's coming down hard on them. And so, it's hard to...it's hard for people to be like, 'yeah, white pride and white privilege when I'm barely finding food over the weekend, too" (Julia*).

These tensions have potentially made it difficult for teachers or the school district to change any curricula. When asked about any changes she would make to have the program be more culturally relatable to her students, she shared,

> "That is a slippery slope, too. I feel like you could, you know, especially in this area, if we as teachers said, 'we have this new, we've made these new revisions'...it's such an angry, divisive time. I think if we came up to this culture, and it came out that we were revising the SEL to include more cultural relevancy and all this stuff, I think it would just become attacked, and it would be front page news, right? Whether it was good or not" (Julia*).

With the understanding that she is a teacher in a rural area, she also expressed difficulty around sharing different perspectives in the school's community:

> "Politically, people, you know, no one wants to feel like they're being targeted for something they didn't do. And at the same... so, like, that would be the white side argument in this area. Like, 'I didn't do those horrible things that were done in history.' The Native American people just want to be heard and given equal rights and freedoms and in the school, in their jobs, all the things and they've been...they're coming from a perspective of never having that for the last, however many years. 200, 300 years, 400. So, you know, they're both...both groups are coming from these different perspectives. And I really think both just maybe want the best but yeah...I would say, white people to just admit that there is a [different] perspective. It's not... it's that Native American people had this really horrible traumatic past, and the more you learn about, like, generational trauma, and all of that icky-ness you can see how that trickles down and it just continues on….But so when we bring in a cultural, cultural ed speaker, for instance, it quickly turns into a 'you versus me' type of... and there's always lots of talk after the meeting, like, 'Well, yeah, but that doesn't really work' or 'that's idealistic,' or, you know, 'there's always this'... or there's people on the other side, 'yeah, we need to do better about that right now, and we need to change all these things.' And then other people are like, 'we can't change these

> things, because that's impossible,' you know, so we've had lots of cultural ed type classes. Or speakers that come in or are like workshops. I don't know that it's made a difference?" (Julia*).

Although the cultural context of the schools' community is only from the perspective of two teachers, they highlight well-recognized phenomena. First, Julia finds it important to teach honest history to the best of her ability because she experienced the failings of U.S. schools' social studies curricula. In particular, historical and present issues of race are briefly–often inaccurately--covered or excluded completely from social studies curricula (Gay, 2002; Ladson-Billings, 2003). Next, Julia highlighted how the intersection of class and race makes it difficult for the rural, white community in the area to recognize they hold a racially privileged identity. This situation is described by scholars as an issue stemming from the prevalence and perpetuation of whiteness in society, and how whiteness fails to account for intersecting identities (Blacksher & Valles, 2021). In part, whiteness is a politically and socially constructed concept where light-skinned individuals, often from European descent, are held as the standard comparison group of all other groups (Calgary Anti Racism-Education, 2021). Thus, when white communities, often living in rural, impoverished communities, are not meeting the ever-changing expectations of "whiteness," (e.g. people get far by working hard), blame can be placed on themselves or on Black and Brown communities (Blacksher and Valles, 2021). Blacksher and Valles (2021) call for a need to discuss the intersections of class and race in a way that will benefit all communities. Lastly, the resistance to change initiatives, particularly those that relate to race or culture have been recognized in other schools with educators facing barriers such as staff attitudes and beliefs, parental expectations, and desire to keep the status quo (Theoharis, 2008).

Teacher CRP Training. Teachers' training was limited to a few different experiences. Some mentioned potentially having a master's course that focused on CRP.

Others mentioned attending professional development sessions with a culture component as a requirement for re-licensure. Several teachers remembered having a training session with an Indigenous speaker. One teacher shared what that experience was like. She stated,

> "We've had people come in and talk to us on Native American culture or Native American history and ways to engage families…So we had Dr. [speaker name] here, for one hour, I think and he kind of just gave the history on Native American tribes in [study state] and then talked a little bit about Ojibwe culture, and then language and then kind of just showed us, because they're adding Native American standards into [study state], social studies and whatever….He gave us some resources, like, 'oh, the [city in the same state] district, like, here are a lot of their resources because they've done a lot of work and it's ready for you'. So, he kind of gave us some resources that we can use to teach that. And I think that's all that this school has done. My last school had him. I think we had three full days of PD with him. And it was more on the historical side of just learning about history" (Natalie*).

Further, teachers discussed training on students from low-economic status backgrounds through book studies on poverty. One teacher also described how some training sessions focused on economic status described the focal points as, "How does poverty shape people?' 'What does that look like for the students who are facing poverty?" (Rebecca). Lastly, another teacher had several, varied experiences with CRP training.

> "I've been a part of lots of different types [of trainings]. I've been a part of some where it's just covering your bias, you know, trying to get the staff to reflect on their own bias, and then maybe even put into words, what their core beliefs are so that you can see how that might translate to the way you teach workshops like that. I've been a part of workshops where we all sit around and do a sage ceremony and talk about the way things might go at home for a Native American student. I've had really

> informational ones where it's just a speaker up front, and he's just showing us facts, like, 'this is how many percent of your Native American students are living in poverty. This is the percent that graduates. This is the percent that's living that doesn't have food over the summer'" (Julia*).

Broadly, these teachers' training in CRP was limited to singular professional development days at their school. Although the teachers mentioned having CRP training was a requirement for re-accreditation, almost none of the teachers mentioned getting training in their teacher education programs. This is not surprising given specific training surrounding issues of race, ethnicity, and culture teachers are often missing in their university training programs (Allen et al., 2017; Hayes & Juarez, 2012; Milner & Laughter, 2011). Further, the informational training as described by Julia demonstrates a deficit-focused narrative regarding statistics of student groups in poverty. A deficit-focus of students' skills centers students' weaknesses, gaps in knowledge, or lack of home support (Kwok et al., 2022). This narrative is common within preservice teacher's beliefs regarding the implementation of CRP where learning factors contain a deficit focus—indeed, preservice teachers' frequently associated low socioeconomic to low vocabulary acquisition (Kwok et al., 2022). This deficit-focused narrative, especially for minoritized students, negatively limits students' sense of educational opportunity (Shapiro, 2014). Table 15 details the types of CRP training teachers reported.

Teachers' Cultural Backgrounds. Next, the interview responses provided an understanding of how the teachers described their own cultures, which allowed for insight into teachers' cultural perceptions of BGP. When asked about their cultures, some teachers had difficulty with how they might answer the question. One teacher started by saying, "This is a really tough question because I feel like culture just 'is'. Like, I really don't think about it, right?" (Laura). Although some mentioned race, religion, and language, most

teacher responses regarding their cultural backgrounds were more expansive when discussing the personal values aspect of culture. This is fitting as shared values build culture (Smolicz, 1981). Some common values included the importance of centering family, hard work, and respect. Respect for one teacher was defined as, "Respect means to me treating everyone fairly, um, to not look down on anyone else to make sure that they're included. And just being kind and open with them" (Rebecca*). When discussing respect, another teacher shared,

> "I think being respectful, like, respecting authority, like, that's important to me. That's, like, I respect the people in authority over me. And so, like, working together, collaboration, like, that stuff that is really important to me. Teamwork, respecting differences, whether you agree with somebody or not like everybody deserves to be respected" (Natalie*).

In general, teachers felt their cultural backgrounds were reflected in the BGP program. Teachers seemed agreeable with the values taught. With the perpetuation of whiteness as a norm for many social situations, one teacher said, "I didn't even notice anything, like, different than me culturally, which probably means it was just like everyone here" (Julia). Specific values that a teacher thought represented core topics and values in BGP included, "treating people where you want to be treated, respecting yourself, and others listening, respecting authority" (Laura). Another teacher shared how she did not believe her personal identities were necessarily the reason she agreed with the curriculum. She noted,

> "Obviously, I don't talk about my religion and my own personal faith and beliefs because I don't find that to be appropriate for school. But [BGP is] still aligned, like, being honest is...whether that's a religious thing or not, I think that's just part of being a good person. And so, …they're still truths that I believe in that are found in

> being good people, being honest and being respectful and setting a good example and not stealing and not lying. All those kinds of things that I believe in, that it aligns with me but it's more of, 'this is how to be a good person' versus, 'this is my personal religion'" (Natalie*).

Despite this teacher's thought that her religion could be separated from her belief of what it means to "be a good person," the perpetuation of whiteness means that many things that are held as the standard, for example, what it means to be a good person, often operate from an unconscious standpoint because of how ingrained whiteness is in society (Calgary Anti Racism Education, 2021).

Interestingly, most teachers tended to highlight the differences between their cultural backgrounds and those of the students. Importantly, teachers' cultural values seemed to differ from their perception of their students' cultural values. From a teacher who grew up with value of doing homework work before play, she stated,

> "There's, you know, 50% of my parents…they do the homework with their kids every night. And then there's the 50% that don't and oftentimes, it's the 50% that needed the most so…And it used to be more like 70/30 and 80/20. And it's becoming more 50/50 now, like, less and less responsible, less and less, you know, structure in their life" (Ava*).

Another teacher who grew up in a family who valued high-structure and authority described differences between her and her students' family lives.

> "I think, for one, I have a very different family life than a lot of my kids. Like, I grew up in a two-parent, very sound-structured home, huge on routines, huge on family. Like, high level of, 'you follow directions' and 'do what you need to do.' And a lot of my kids just don't have that same family structure at home, and they don't have the

> same support, and they don't have the same expectations for the way that they treat their friends or their, like, the way to talk to parents. I'm, like, astonished" (Natalie*).

When asked about how their cultures compared to their students, teachers highlighted the influence of technology on students' lives. One teacher made it clear how her childhood and her students' childhoods differ. She emphasized how family values today look different than her family values did with the influence of technology.

> "Technology has become more important in people's lives, and they spend more time on their technology and less time as family. Yeah, I think my own children, we always had dinner together all the way through their whole life and even in high school, you know, we've always had dinner together and many of these kids you don't see that. They're, they're on the run. Their lives are so busy that they're so scheduled, that, that sit down quality time together as a family to talk about your day. It just doesn't happen with a lot of them" (Susan*).

This teacher's observations regarding technology use and the differences in generations have been noted in the literature. Increasing, novel technology use has rapidly changed enduring facets of culture such as how people communicate, what they believe, and how they act (Combi, 2016). When discussing economic status, a teacher who stated she came from a middle-class background found that she differed in family spending values.

> "Well, even different things about thinking in terms of what they would spend when they do have money, what they prioritize and what they would spend their money on, versus, things that I would necessarily prioritize. It's for, like, entertainment always seems to be a big one. Some of our kids that they seem they don't have anything, but they come to school and they have cell phones, some of those kinds of things. It's definitely different, but that's what they value, and that's what they would see important. So some of that, you know, just how they... their home lives are

> just very different in how they're structured and maybe even how they're treated at home and some of those things are different" (Christina*).

Christina described herself as coming from a middle-class family. Her interview quote above indicates potential cultural value differences between economic classes, in terms of spending. This belief has shown up in various narratives surrounding poverty and how people believe individuals in low-income communities should or should not spend their money (Miller et al., 2021). See Table 16 for a more detailed breakdown of the cultural values that teachers mentioned. With this background understanding of the locale context, teachers' training, and teachers' cultural backgrounds, a single theme regarding teachers' perceptions of BGP was identified.

Theme 4: Teachers' Believed BGP is Culturally Universal

The teacher's interview responses then illuminated their thoughts on the cultural relevance of BGP. When asked about how BGP was made for their students, teachers tended to respond with how BGP is both fitting for their students and for all students. One teacher succinctly stated the necessity of the skills. "It's teaching them things that, um, that they should know. That they need to know" (Ava). Other teachers described the diversity of the photos and different scenarios as the reason why BGG is designed for all students.

> "It covers pretty much all of our students...they got different pictures, and they've got different, you know, ethnicities, and people are represented in the program. The scenarios…they'll have ones at school, and they're relevant to everybody, just because that's a safe place. And that's, you know, and then they've got the different examples at home. They use different names and different…examples of… I think it was one where that kid was eating something that was maybe, a Korean…or something different, and our kids are just kind of, like, 'well, what on earth is that?' But it was just a good example to show our kids that aren't exposed to a lot of

> things…it's a curriculum that's designed for all students. So, I think it's, it's built pretty well" (Christina*).

Similarly, another teacher detailed,

> "Even the character name choices, like, the characters' names, and the scenarios aren't, like,
>
> 'Mary, John.' …A couple weeks ago, there was 'Juan' and there's other names from other cultures. And that might be one way that their culture could be reflected in the Be Good People is just names that they recognize that isn't necessarily you know…. I guess I feel like it's made for everybody. I don't feel there's anybody that's excluded…I don't feel like it's biased in one religion or in one culture. It's very much applicable to everybody. Like no matter where you're from, no matter what language or what your home life is like, or what your education background is like. It's universal truths, really, that I think are in a part of Be Good People. So, I don't know…. I think it's a good fit for all kids" (Natalie*).

Although the teachers noted these as strengths of the program, the CRE Scorecard results contrasted with these findings. BGP was rated culturally insufficient largely due to the ambiguity of the characters. Thus, despite the diverse depictions, little meaning was attached to the images given their vague nature. These actions lend to a level of tokenism, where non-white characters and names are used at a surface level, rather than engaging richly in their identities. Tokenism in literature can leave audiences feeling the chaotic lack of intentionality behind diverse characters (Podoshen et al., 2021). Further, a teacher made the assertion that children don't notice racial differences in first grade, when asked if there was anything she would change about BGP to culturally fit the needs of her students better,

> "There really isn't. Other than there is no talk of how to discuss different races or ethnicities, which in first grade, they don't even see that. Honestly, kids are just open

> and kind until they're not, but they're, they're not unkind, like 'your skin's Brown and my skin's white.' Like, they don't, they really don't think that way yet. Which I'm very grateful for because it makes it, makes it easier sometimes but, but I do think, you know, like, there could be a lesson just like, 'you might have someone who doesn't celebrate Christmas someday in your life, like how are you going to handle that?' You know? and it may take them till they're in high school before they meet someone who doesn't celebrate Christmas, but at least if we've talked about it every year since they were in kindergarten, they'll be like, 'Oh yeah, this happens,' you know?" (Julia*).

Contrary to Julia's beliefs, several studies have found young children can identify different races. For example, nine-month-old infants were able to form categories based on race (Anzures et al., 2010). Children as young as 3.5 years old use race as cues for wealth (Mandalaywala et al., 2020). Julia's perception of when children start noticing race is not uncommon. Indeed, one study found that parents predicted children's abilities to process race was approximately four-and-a-half years the prediction of scientific literature (i.e., 36 months; Sullivan et al., 2020). This misunderstanding of when children could process race was also a strong predictor in adults delaying conversations about race (Sullivan et al., 2020). Moreover, despite the teachers' perceptions that BGP is for all students given the diversity of depicted students in the modules, names in the scenarios, and belief that young children do not notice race, teachers still believed there could be changes made. In the following research question, one of the changes that teachers believed could be made was including culture-specific scenarios.

Research Question 7: Changes Teachers Would Make (If Any)

Two themes were identified for the types of changes teachers would make to the program. The first theme contradicted the idea of universality of the BGP program.

Theme 5: Teachers Would Add More Culturally Relevant Scenarios

Although most teachers believed BGP was suitable for all students, one teacher did not share similar sentiments. Regarding the scenarios she stated,

> "I definitely don't think they're very culturally relevant, relevant to anyone that isn't white. You know, like I think there were pictures…You know, the kids had color in the pictures or whatever. But like I said, if I had a student that celebrated Kwanza in the classroom, which I don't, and I have never have, but I think you know, that could be a really good lesson and maybe it could be brought up a couple more times, or even the fact that Native American children typically don't make eye contact because it's, it's considered disrespectful and rude. But, you know, in a white family, that's something so simple, but in a first grader, they might not understand that so you could talk about that" (Julia*).

Despite teachers believing BGP is suitable for all students, they still identified more culture specific scenarios that could be added including racial and ethnic specific cultures, holidays, and family structures. For example, one teacher said,

> "Maybe a few more examples from different cultures that they haven't experienced yet. I know there's a few in there. Native Americans aren't really touched on and this area has a very huge population of Native Americans. So, some more examples of that. Some more examples of Indian culture, which we don't really have a big population here but a lot of people don't touch on that culture very well. Yeah. There are cultures… they... aren't seeing [in] their everyday life but get them involved and thinking about, 'Oh, other people might not celebrate Christmas because everybody here celebrates Christmas.' …So, getting them open to that possibility of Kwanzaa and Hanukkah and that" (Rebecca*).

Another teacher wanted BGP to highlight the intersectionality of different cultural aspects.

> "So, Native Americans, more than that, their customs or differences, open dialogue. And even within the same ethnic backgrounds, like married versus divorced, parents, like that's a different culture, you know? Students that don't have their parents and are living with grandma and grandpa" (Laura).

Table 17 provides more detailed information about the types of culturally relevant scenarios that teachers would add to the BGP program.

Theme 6: Teachers Would Increase the Developmental Appropriateness of BGP

Another main theme identified for this research question was teachers' desire for BGP to be more developmentally appropriate. One teacher mentioned,

> "I would add more things to increase student engagement, just little activities to get the kids kind of moving a little more rather than sitting…content that's probably a little more friendly for six and seven-year-olds. Sometimes, I felt like some of it was a little deeper than they were capable of understanding…the younger kids with their being so centered on themselves…If there's an activity that I think of that would be more hands on for the kids I try to put that in on one of the other days. If it's kind of relevant to what we're doing. So, I mean, I tried to do that stuff to make it more hands on for the kids" (Susan).

One teacher provided a suggestion to simplify the steps in each of the scenarios.

> "I think [BGP is] easy to understand. The language content is probably my biggest issue with it. The scenarios are good, you know, the digital accessibility that's all good. It's too many words on the board, really is the problem…I think less words, so you know, whatever the rule is, whatever that main point is, if it's like, 'stop and think about what..., how you're acting,' we can simplify down to the word 'stop.' Or if it's, you know, 'think of a kind of apology,' and whenever it needs to be like 'apologize,' you know, one, one or two words that you can really have them latch on

> to. And then a similar protocol for every, you know, I think, like, "stop and think about it" is a pretty similar protocol for every situation you're in, or most situations, obviously not emergencies and stuff" (Julia*).

A teacher also reduced the length of scenarios near the end of the school year to allow for creativity within her students.

> "Sometimes, I shortened down the scenarios at the end, and then I have kids come up with their own scenarios. Or yesterday…this week is sportsmanship and so they had to get in small groups and make a skit demonstrating sportsmanship. So, it's a little bit more like they come up with a scenarios and they're acting it out, and it's more geared towards, like, their real life school scenarios…the scenarios get to be a little bit like long, and I'd almost rather have, like, two or three scenarios and, like, go into depth with them and, like, really hash them out and do, like, perspective taking and stuff like that with them versus, like, having to go through, like, six or seven, like, short little things, especially because fourth grade, like they know the answers, like, everybody knows the answers" (Natalie). Please note–every BGP scenario in the same module has the same steps for problem solving, which is likely what this teacher is referencing when she says, "everybody knows the answers.

See Table 18 for more specific types of developmentally appropriate scenarios teachers would add.

Discussion

This initial investigation into the BGP program revealed mixed findings. When considering the limited efficacy testing, BGP demonstrated promise for continued implementation in this school district given the increase in teacher ratings of students' social and emotional skills, increase in academic grades for two of the four grade levels, and high, quantitative acceptability and feasibility ratings. Study findings for improvement areas

of the program include increasing the cultural relevance and developmental appropriateness of the program.

The first hypothesis was partially supported, where teachers' perceptions of students' self-awareness, self-management, and relationship skills increased from T1 to T2 of this study. However, their perceptions of students' social awareness and responsible decision-making skills had not. The teachers not having perceived an increase in all SEL skill areas in the short period of the study is consistent with a developmental lens of SEL growth; the skills measured by the SSIS-SEL screener take years to develop and refine as students age (Denham, 2018). Laura described SEL at the elementary level to be more building a toolbox of skills to use in the future. Indeed, a meta-analysis including over 97,000 students indicated that students who were involved in universal SEL programming ended up with higher levels of social and behavioral outcomes such as peer acceptance, peer relationships, and attendance up to 18 years later compared to students who were not involved in universal SEL programming (Taylor et al., 2017). Moreover, what mastery looks like in each CASEL domain is not defined by developmental level in the SSIS-SEL screener; thus, the screener could benefit from adding developmental considerations of what mastery of each skill looks like at different grade levels.

The second hypothesis was also partially supported. As hypothesized, students in first, second and third grade had a statistically significant increase in academic achievement. However, fourth grade did not. The improved academic performance among the younger grade levels is consistent with the finding of improved academic grades across several SEL studies (Cipriano et al., 2023). The lack of a statistically significant increase in fourth grade academic grades could potentially be attributed to a ceiling effect where most students were already receiving the highest grades possible at T1. For the mySAEBRS and SAEBRS-TRS outcomes, students' and teachers' ratings did not statistically significantly

change from T1 to T2. This could be a result of conducting Wilcoxon signed rank tests on students total score values, as opposed to the categories (i.e., "no-risk," "some-risk," "low-risk") in which those values. Thus, an exploratory chi-square test was conducted to better understand students' risk levels. Given that most students were in the "no-risk" category for both the mySAEBRS and SAEBRS-TRS at T1, there was potentially a ceiling effect where students were already at the lowest risk level possible.

A large benefit of the BGP program is its practicality levels. With most teachers rating and describing the practicality of the program positively, BGP holds promise as an SEL program that could overcome some of the common barriers of implementing SEL. One qualitative study found that teachers' limited time to plan lessons for the school-developed SEL program was one of the largest barriers to implementation (Martinez, 2016). Another study of 263 teachers identified the top three barriers to SEL implementation being 1) time available to teach the lessons, 2) time available to prepare for teaching the lessons, and 3) resources to purchase an SEL program (Buchanan et al., 2009). Beyond BGP being a freely accessible curriculum, teachers in this study found the feasibility of implementation to be a strength of the program. Many noted the minimal preparation they needed to plan lessons and implement the program. The practicality of the program also comes with the understanding that the school district prioritized SEL. This is foundational, as teachers' support of SEL implementation can be influenced by their school's culture of support surrounding SEL (Brackett et al., 2011).

The program was deemed acceptable to teachers as rated by the URP-IR. In general, teachers rated most items as "agree" or "strongly agree." The lowest rated item was "This intervention is a good way to handle the child's behavior problem" and the highest rated item was, "I would be committed to carrying out this intervention." The lowest rated item was, on average, rated at a "slightly disagree" level. The lowest rating could be accounted for

by the problem behaviors that elementary school teachers face. When considering the broad scope of "behavior problems" at the elementary level, SEL may not necessarily provide an immediate solution. Often, problem behaviors at the elementary level are when students engage in disruption, disrespect, and fighting (Algozzine et al., 2008). Evidence based strategies for teachers for these types of problem behaviors at the classroom level can include praise and effective commands (Owens et al., 2017); whereas SEL can often target broad behaviors (e.g., "responsible decision-making") and is a continual, developmental process (Casel, n.d.-a). However, teachers rated their commitment to SEL as one of the highest items in the acceptability subscale of the URP-IR. This provides high promise for future studies of BGP as teacher commitment is a main factor for successful SEL implementation (Brackett et al., 2011).

The CRE Curriculum scorecard and teacher interviews yielded conflicting results. With the ambiguity of characters presented in the BGP scenarios, lack of encouragement to think critically about inequities, and little guidance on how to engage diverse learners, the ratings on the CRE Curriculum scorecard places BGP in the culturally insufficient category across. Yet, in this sample of white teachers serving primarily white students, they expressed that the program was universally applicable for all students, despite highlighting the need for more culturally relevant scenarios. Although large meta-analyses have shown universal SEL programs to be positive in supporting student outcomes, researchers note that much more work needs to be done to better understand how universal SEL programs support minoritized students (Cipriano et al., 2023).

Limitations

Despite some promising findings, this study is not without its limitations. The first is the nature of a feasibility study. Limited efficacy testing is often conducted with a convenience sample and designed to test intermediate rather than final outcomes (Bowen

et al., 2009). Therefore, no causal relationships could be established in this study. Next, the study was conducted with a small sample size of teachers who came from a largely homogenous background. The teachers' backgrounds were also similar to that of their students. Therefore, the results of this study may not generalize to different settings (e.g. suburban or urban) or diverse populations. Additionally, the change in academic performance was based on different academic standards from T1 to T2 for the students as no pre-established pre- and post-intervention academic screening measures or progress monitoring was leveraged. Lastly, the CRE Curriculum Scorecard results were not scored with any members of the community, which is not recommended by the guidelines for scoring. The results of the scorecard should instead be interpreted with extreme caution because the individuals scoring did not represent multiple roles (e.g., teachers, students, families, non-experts, experts, etc.), as suggested by the scorecard guidelines.

Implications for Practice

From this study, authors have gleaned that BGP is an SEL program that could be implemented with fidelity and has high practicality perceptions from teachers. Analyses of the acceptability of the program indicate that teachers are committed to the program. Concordantly, in the interviews, teachers believed that the skills taught in BGP are important for students to learn. Yet, there were a few areas for improvement that the teachers identified across different measures in this study. One of the main areas for improvement was improving the developmental level of the materials. Thus, for educators who wish to leverage BGP curriculum in the future, they may consider the developmental level at which the scenarios are written. Teachers in this study noted the difficulty for students to understand various SEL concepts and having to adapt scenarios to fit the developmental level of their students. If BGP continues to keep modules together by grade band (i.e., lessons K-3 instead of having modules specifically for individual grade levels),

educators could take a few steps to ensure the modules are developmentally appropriate for their students. First, educators could select modules that appear developmentally appropriate in language and examples for their students and create a schedule of lessons to teach composed of developmentally appropriate modules. Another method for increasing the developmental appropriateness of this program could be to align module selection with state SEL standards. Although 27 states have standalone SEL standards, few have ensured that these benchmarks are developmentally appropriate (Eklund et al., 2019). For example, the state standards that BGP is aligned with are listed in grade bands for each of the five core competency areas of SEL (Minnesota Department of Education; MDE, n.d.). This could be an area for educational standards developers to ensure state standards are more specifically aligned with developmental levels.

Next, this study gained important understanding about implications for practice regarding the cultural relevance of the program. Although the teachers in the study found BGP to be applicable universally, the context in which this study took place must be accounted for with this interpretation. In particular, we emphasize the backdrop of this study—a rural-based school district with a predominantly white student body. The racial identity of teachers then also matched that of the majority of the student body. In the United States, whiteness exists as the "normal" culture, thus, customs, cultures, and beliefs of other cultures are compared to those of white racialized identity (National Museum of African American History and Culture, n.d.). In this study's location, Julia shared how any change to the BGP program to include the discussion of culture could potentially cause an uproar. This falls in line with many, current discussions around SEL in the United States.

Several states have proposed bills to ban SEL in schools, due to concerns that SEL distracts from academics and instead pushes left-wing propaganda (Abrams, 2023). Accordingly, the culturally insufficient score we rated BGP would mean that the program

would likely be better received in areas where there is resistance to ideas related to equity and social justice. Yet, we cannot underscore enough the importance of CRP when designing these universal programs because culturally diverse students can be found everywhere. With the growing diversity of schools in the country (NCES; National Center for Education Statistics, 2022), culturally relevant education is important to the growth of students who do not fit into dominant culture (Bennouna et al., 2021). One way for BGP to push toward supporting diverse populations is for developers to examine the CRE Curriculum scorecard and push toward the idea of transformative SEL, where SEL competencies begin to align with constructs for social change (Jagers et al., 2019). Jagers et al. (2019) provide example competencies that align with personally responsible, participatory and transformative SEL. For example, responsible decision-making on the personal responsible level looks like learning about personal well-being whereas at the transformative level, it looks like focusing on collective well-being.

Implications for Future Research

The authors conducted limited efficacy testing with a focus on the practicality and acceptability of the BGP program. These aspects of the feasibility study, with an additional focus on cultural relevance of the program, provide several avenues for future research. First, future methodologies could employ quasi-experimental and experimental designs and leveraging sampling strategies and conditions that could allow for causal interpretations of data. Additionally, a longitudinal study could be conducted to have a better understanding of the program's impacts on student academic and behavioral outcomes over time. This study design could allow for understanding more nuanced growth in SEL skills using formal academic, social, and emotional screening and progress monitoring measures. However, given the, often, unpredictable nature of applied studies in schools, future researchers could conduct feasibility studies of BGP in suburban and urban schools with diverse populations

to better understand not only the feasibility of BGP but the cultural relevance of the program in those areas. Increasing our understanding of how BGP functions in diverse settings can allow for a more nuanced understanding of how this universal SEL program functions outside of rural, predominantly white spaces. BGP is also an expansive program in that there are several facets that were not examined in this study such as the usage of extension activities or the academic integration aspect of the program. Some researchers have also called for SEL to be more integrated into daily schedules rather than being standalone lessons (Bailey et al., 2019). With the resources available on BGPs website regarding academic integration (Be Good People, n.d.), this could be a promising program to explore.

Conclusion

This initial investigation into the BGP program revealed that the program demonstrates some promise when implemented in a rural school-based setting. Although the practicality of this SEL program was rated highly, two major areas for improvement include increasing the developmental appropriateness of the program and expanding the cultural relevance of each module to broader student populations. When thinking about future uses of this or any SEL program, educators and researchers should consider how to ensure the cultural relevance of the program, so all students have the opportunity to reap immediate, gradual, and meaningful benefits.

Chapter 4

Synthesis and General Discussion

The current state of our education system includes an increasingly diverse student body and rising costs for SEL programs. Thus, this multi-study dissertation explored the landscape of social, emotional, behavioral, and academic student outcomes in the context of an increasingly diverse student body and locally developed social emotional learning programs. The first study was a systematic review of a relatively few numbers of studies that investigated the impact of cultural adaptations on student SEL outcomes. Although the effectiveness of these outcomes was mixed, many of the studies found high acceptability and feasibility for the cultural adaptations that were made. The feasibility study of BGP in the second study of this dissertation revealed that the BGP SEL program demonstrated promise as there was an increase in academic grades and SEL skills across the short-term study. However, despite the "universality" claims found in the teacher interviews, the CRE Curriculum scorecard and teachers highlighted the necessity for more cultural relevance in the program. Specifically, teachers called for more practice scenarios that would be relevant for the Indigenous youth of the area, along with including more diverse family structures. Together, the studies demonstrated the need for more research on the intersection of SEL programs and cultural relevance, whether that be in adapted SEL programs or a locally developed program.

Implications for Practice

Although implementing SEL programs in schools have demonstrated wide social, emotional, behavioral, and academic benefits for students (Cipriano et al., 2023), there are many considerations for educators when choosing to implement an SEL program in a school. The dominant culture in the U.S. positions values associated with whiteness and the middle class in SEL programs to be considered the norm, leaving students from lower-

income, racially and ethnically minoritized backgrounds to face negative biases and higher stress levels (Jagers et al., 2018). Thus, educators who work in diverse school settings where students may not relate to white, middle-class norms should take extra caution when implementing SEL programs to ensure they are culturally relevant for their student population. From the first study, a feasible and acceptable method of introducing cultural relevance in SEL programs is including community and student voice for the types of adaptations to make.

Further, the second study indicated a need for not taking "universality" at face value. There are a few steps educational practitioners can take toward ensuring SEL programs are culturally relevant for students. The first is to examine one's personal biases towards the values that SEL programs teach. In the professional educational body, most teachers are white women who no longer represent the racial and ethnic backgrounds of the majority of U.S. students (Schaeffer, 2021). As Julia stated, she didn't notice anything different culturally about the BGP curriculum from herself. When educators find themselves not questioning educational practices, this could be a sign of complacency, and following the narrative of the dominant culture (Tevis & Foste, 2023). Thus, frequently critiquing one's biases on what culturally is and isn't included could be beneficial in working toward cultural relevance of these, supposedly, universal SEL programs. Second, the change toward more equity-centered, culturally relevant curricula cannot be a minor goal of a district. To have more culturally relevant curricula, educators must recognize the need for culturally relevant curricula (Sharma & Christ, 2017). Systems change for the types of SEL programs that are both created and implemented will require dedicated time and patience as each context in which SEL programs are implemented can vary by location, population, and SEL acceptance.

Critically analyzing SEL programs could be practiced by educators with the CRE Curriculum Scorecard. Although the scorecard was not initially developed for SEL programs

nor does it have psychometric evidence, the collaborative nature of the scorecard could allow for a baseline understanding of how culturally relevant an SEL program is. However, given the attacks SEL is under (e.g., "SEL is being used to teach white children to hate themselves;" Anderson, 2022), failing to be "racially neutral" in some communities could leave SEL programs vulnerable to removal with several states recently introducing bills to ban SEL from their schools (Bader, 2023). Consequently, the need for research to demonstrate the importance of cultural relevance within SEL programs and inform the public on this intersection is important.

Implications for Research

In general, there is much more research necessary at the intersections of cultural relevancy and SEL programs. The systematic review revealed that only a handful of studies have examined the impacts of cultural adaptations on SEL programs. Further, these studies were conducted with limited student populations, thus the generalizability of cultural adaptations to SEL programs in other minoritized populations is not currently known. Similarly, in terms of locally developed SEL programs, future researchers could examine cultural relevance of the SEL program's content with populations for which the program was not developed. Additionally, although core components of evidence-based SEL programs have been identified (i.e., self-awareness, self-management, social awareness, relationship skills, responsible decision-making; Wigelsworth et al., 2021), integrating these competencies in a way that centers culturally relevant instruction is an area with little research. Despite extant research discussing how to prepare educators for culturally relevant teaching through building their recognition of personal and other socialized identities (Ryan, 2002), this point was not uncovered in either the systematic review or feasibility study, although there was opportunity to do so, highlighting the need for future research to explore how to make SEL programs more culturally relevant. Future

researchers could also focus on improving the following aspects of culturally relevant SEL: (1) Identifying for whom and what reason are cultural adaptations being made to an SEL program; (2) What student, family, and community input is being included when deciding if an SEL program is culturally relevant for a school's student population; (3) How future program-developers can incorporate an equity-centered approach to SEL programs so that SEL programs don't rest on the idea of a "universal" approach.

Conclusion

Understanding student outcomes related to SEL programs in the U.S. requires context for the cultural relevance of the SEL program to the students who are learning from the program. Both studies have underscored the lack of research in the areas of cultural relevance in SEL. For students, and broader society to begin reaping the benefits of SEL beyond social, emotional, behavioral and academic outcomes in schools, future research and practice necessitates centering equity in SEL programs.

Closing Reflection

When I reflect on my educational experiences from kindergarten through graduate school, I think about how my greatest moments for learning were filled with compassion, patience, understanding, and acceptance from peers, instructors, and myself. When I think about the young version of my siblings and myself who navigated a school system that wasn't built for children who looked, spoke, and understood the world like us, I feel overwhelmed—and quite frankly, sad. These emotions do not only stem from our personal experiences, rather, it's the understanding that so many students' cultural backgrounds are seen as negatives: odd, bad, or foreign. Instead, they could be viewed as strengths in building a community. Valuing cultural diversity is not only a strength, but supporting cultural relevance in education is also a necessity for educational equity. I recognize that the fight for culturally relevant education sits at the intersection of broad and beautiful

identities. For me, the benefits that schools could reap from incorporating SEL into their students' days goes far beyond the classroom walls. Not only do I imagine a future where students feel safe, comfortable, and seen at school, but I envision a future where students recognize injustices beyond the lunchroom, playground, and classroom—and they not only recognize but act upon injustices they witness. Some teachers in this study noted that the benefits of SEL may not be instant; they hoped that by the time students are young adults, they would have the skills to work with people from different backgrounds. It is with the understanding that SEL is planting seeds today in hopes of gaining a bountiful harvest, that I write this dissertation. I aim for other researchers, educators, and myself to remember that slowing down and framing SEL from an equity standpoint—to ensure we have a bountiful harvest for everyone in our communities—is of great importance. I believe that communities that come together and ask for SEL to be integrated into their schools hold great power in generating a more compassionate, patient, understanding, and accepting world.

Figure 1

PRISMA Flowchart

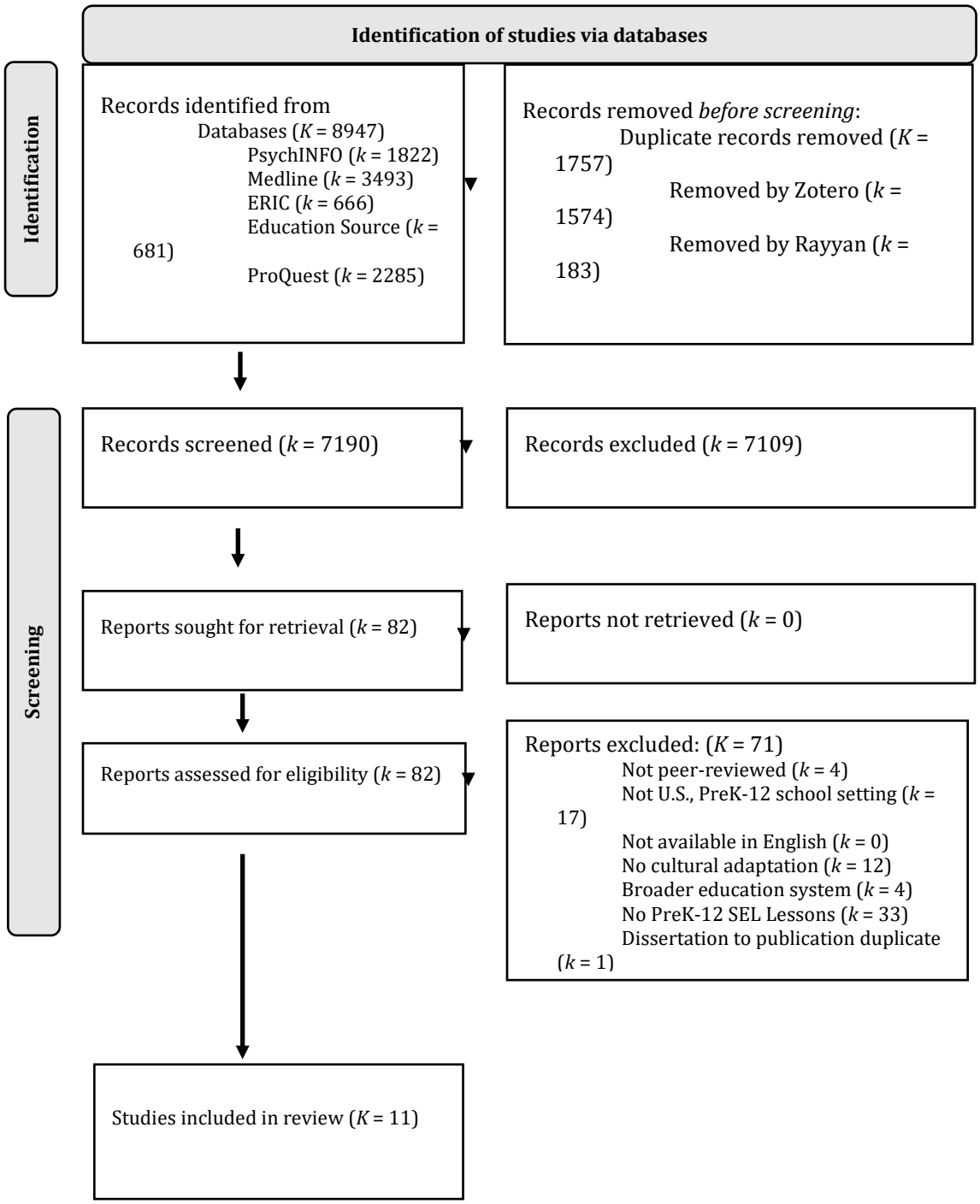


Table 1

Study Aim(s)

Authors (Year)	Research Question(s) or Purpose(s)
Castro-Olivo and Merrell (2012)	1) Assess the program's feasibility of implementation in a regular classroom setting 2) Assess the program's effectiveness at teaching SEL skills 3) Assess the program's social validity and participant satisfaction
Castro-Olivo (2014)	1) To what extent do students who participate in the culturally adapted Jóvenes Fuertes SEL program report higher levels of SEL knowledge and SEL resiliency in comparison to those in the waitlist control group at post intervention assessments? 2) To what extent do participants perceive the culturally adapted Jóvenes Fuertes SEL program as a socially valid intervention
Castro-Olivo et al. (2022)	1) To what extent the proposed adapted factor structure of the BERS-2 adequately measures each of the 5 SEL competencies (i.e. relationship skills, self-awareness, responsible decision-making, social awareness, and self-management) in this sample? 2) To what extent does a culturally-adapted SEL intervention improve each SEL competency (i.e. relationship skills, self-awareness, responsible decision-making, social awareness, and self-management) for this ELL sample?
Cramer and Castro-Olivo (2015)	1) Is there a significant difference between pre-intervention and post-intervention self-ratings of student resiliency? 2) Does a significant difference exist between post-intervention and follow-up ratings of student resiliency? 3) Do students report a significant difference between pre-intervention and post-intervention ratings of internalizing problem symptoms? 4) Does a significant difference exist between post-intervention and follow-up ratings of student internalizing problem symptoms? 5) To what extent do CLD students find the culturally adapted Strong Teens curriculum socially valid?

Authors (Year)	Research Question(s) or Purpose(s)
Elek et al. (2010)	This study examined the effects of an enhanced version of the Keepin' it REAL curriculum to addresses issues related to the transition faced by youth experiencing acculturation
Graves et al. (2017)	1) Does implementation of Strong Start result in an increase in knowledge of social and emotional skills in urban African American students? 2) Does Strong Start increase the social-emotional assets in urban African American students? 3) Does implementing Strong Start decrease externalizing behavior in urban African American students?
Hecht et al. (2018)	The study examined the relationship between delivery quality and program adaptation.
Hopson (2008)	1) Is an adapted version of Keepin' it REAL able to reduce intentions to accept offers of substances among participating alternative school students? 2) Is an adapted version of Keepin' it REAL able to reduce substance use among participating alternative school students?
Kurtz (2020)	1) What are the opinions of involved parties of the adapted Strong Start and Strong Kids lesson plans for implementation in an urban school? 2) Are there significant differences in outcomes for students who receive the adapted Strong Start and Strong Kids intervention class wide versus in a small group, as measured by Negative Affect scores on the universal behavioral health screener: Behavior Intervention Monitoring Assessment System 2 (BIMAS-2)? 3) Are there significant differences in outcomes for students who receive the adapted Strong Start and Strong Kids intervention versus the district mean, as measured by Negative Affect scores 4) Do changes in Negative Affect scores on the BIMAS-2 in response to the adapted intervention maintain from post-intervention to follow up?
Martinez (2011)	This study was a program evaluation that appraised the Spark Program for its potential to produce significant changes in at-risk children's emotional and behavioral outputs in key areas such as anger management, educational aspirations, peer attachment and self-esteem.

Authors (Year)	Research Question(s) or Purpose(s)
Moran et al. (2019)	1) Do students trained in the adapted STAC intervention report an increase in knowledge and confidence to intervene as "defenders?" 2) Do students trained in the culturally adapted STAC intervention use the STAC strategies when they observe bullying? 3) Were there differences in outcomes between White and Hispanic students? 4) What were students' experiences participating in the culturally adapted STAC intervention and using the STAC strategies to intervene in bullying situations?

Note. The author(s)' language was used when possible for the research questions and purposes column. The authors of this systematic review recognize that language is ever-changing. There is currently controversy around potential colonialist roots to the term "stakeholder" (Reed, 2022). Thus, the term "stakeholder" from the Kurtz (2020) dissertation was intentionally changed to "involved parties."

Table 2

Study Setting and Design

Authors (Year)	Setting of school(s) and/or district(s)	Design	
		Analysis Type	Method
Castro-Olivo and Merrell (2012)	Schools (N = 1) East Bay area of Northern California	Within group	Pre-post intervention design
Castro-Olivo (2014)	Schools and Districts (N = 7 schools in N = 3 Districts) Southern California	Between group: intervention vs control	Quasi-experimental intervention/ control pre-post
Castro-Olivo et al. (2022)	Schools and Districts (N = 7 schools in N = 3 Districts) Southern California	Between group: intervention vs control group	Secondary data analysis (original study was quasi-experimental)
Cramer and Castro-Olivo (2015)	Schools (N = 1) Suburban, Southern California	Within group	Quasi-experimental pre- and post-intervention group design
Elek et al. (2010)	Schools (N = 30) Phoenix, Arizona	Between group: 5th grade implementation vs control/ 7th grade implementation only vs control/ 5th and 7th grade implementation vs control	Quasi-experimental (longitudinal)
Graves et al. (2017)	Schools (N = 1) Urban, Mid-Atlantic Region	Between group: intervention vs control	Randomized delayed treatment control

Authors (Year)	Setting of school(s) and/or district(s)	Design	
		Analysis Type	Method
Hecht et al. (2018)	Schools (N = 39) Rural Pennsylvania and Ohio	Between group: non-adapted curriculum vs adapted curriculum vs control	Randomized control trial
Hopson (2008)	Schools (N = 4) Urban (*n* = 1 school) Other schools not specified	Between group: intervention vs control	Mixed methods, quasi-experimental pretest posttest design with a six-week follow-up and qualitative procedures
Kurtz (2020)	Schools (N = 1) Urban, Boston, Massachusetts	Between group: small group vs class wide - **Small group:** one combined second and third grade separate classroom for students with learning disabilities and Autism Spectrum Disorder, one combined fourth and fifth grade separate classroom for students with learning disabilities and Autism Spectrum Disorder, one general education fifth grade classroom, one second grade inclusion classroom - **Classwide**: one kindergarten inclusion classroom, one third grade Sheltered English Immersion classroom, one fourth grade general education classroom	Quasi-experimental group design

Authors (Year)	Setting of school(s) and/or district(s)	Design	
		Analysis Type	Method
Martinez (2011)	Schools (N =2) Bay Area, California	Within group	Secondary analysis
Moran et al. (2019)	Schools (N = 1) Rural, Northwestern, U.S.	Within group	Mixed methods

Table 3

Participant Information

Author (Year)	Total	Grade	Race and/or Ethnicity	Gender or Sex Assigned at Birth
Castro-Olivo and Merrell (2012)	N = 40	High school	Latino, recent immigrant (100%)	Female (50%) Male (50%)
Castro-Olivo (2014)	N = 102 Intervention: *n* = 49 Control: *n* = 53	Middle school and high school	Latino (100%); **Intervention:** Mexican origin (15%) United States Origin (15%) Other origin (3%) **Control:** Mexican origin (25%) United States origin (25%) Other origin (1%)	**Intervention:** Boys (26%) Girls (20%) Did not report (0%) **Control:** Boys (23%) Girls (29%) Did not report (1%)
Castro-Olivo et al. (2022)	N = 102 Intervention: *n* = 49 Control: *n* = 53	Middle school and high school	Latino (100%) **Intervention:** Mexican origin (15%) United States Origin (15%) Other origin (3%) **Control:** Mexican origin (25%) United States origin (25%) Other origin (1%)	**Intervention:** Boys (26%) Girls (20%) Did not report (0%) **Control:** Boys (23%) Girls (29%)

Author (Year)	Total	Grade	Race and/or Ethnicity	Gender or Sex Assigned at Birth
				Did not report (1%)
Cramer and Castro-Olivo (2015)	N = 20	9th and 10th grade	Latino/Hispanic (75%) African American (15%) Caucasian (5%) No response (5%) United States origin (75%) Mexican origin (25%)	Female (20%) Male (20%)
Elek et al. (2010)	N = 1,984	5th grade & 7th grade	Mexican or Mexican American (75%) Other Hispanic background (3%) non-Hispanic white (5%) African American (9%) Native American (3%)	Female (50%) Male (50%)
Graves et al. (2017)	N = 61	Kindergarten, 1st, and 2nd grade	African American (100%)	Male (100%)
Hecht et al. (2018)	N = 2781 Non-adapted: *n* = 664 Low quality: *n* = 329 High quality: *n* = 335 Adapted: *n* = 1022 Low quality: *n* = 590 High quality: *n* = 432 Control: *n* = 1095	7th, 8th, and 9th grade	European American (97%)	Female (49%) Male (51%)

Author (Year)	Total	Grade	Race and/or Ethnicity	Gender or Sex Assigned at Birth
Hopson (2008)	N = 41 Intervention: *n* = 18 Control: *n* = 23	High school	**Intervention:** Black/African American (5.6%) Hispanic (38.9%) white (38.9%) other (16.7%) **Control:** Black/African American (8.7%) Hispanic (47.8%) white (30.4%) other (13%)	**Intervention:** female (61.1%) male (38.9%) **Control :** female (56.5%) male (43.5%)
Kurtz (2020)	N = 43	Kindergarten, 2nd, 3rd, 4th, 5th grade	Hispanic (81.4%) white (18.6%)	Female (51.16%) Male (48.84%)
Martinez (2011)	N = 38	4th grade – 8th grade	Latino/Hispanic (68%) African-American/Black (13%) white/Caucasian (2%) Other (13%)	Female (36.8%) Male (63.2%)
Moran et al. (2019)	N = 63 Qualitative: *n* = 23 (subset of the N = 63)	Elementary school	**Quantitative analysis:** white (55.4%) Hispanic (44.6%) **Qualitative analysis:** white (65%) Hispanic (35%)	**Quantitative analysis** females (57.1%) males (41.3%) other (1.6%) **Qualitative analysis:** female (65%)

Author (Year)	Total	Grade	Race and/or Ethnicity	Gender or Sex Assigned at Birth
				male (35%)

Note. In terms of race/ethnicity and gender or sex assigned at birth, the original author language was kept. However, given the gender-inclusivity of the term "Latine", the Martinez (2011), Castro-Olivo and Merrel (2012), Castro-Olivo (2014), Cramer and Castro-Olivo (2015), and Castro-Olivo et al. (2022) had the term "Latino" replaced with "Latine." The authors of this systematic review also recognize there is still much controversy around terms to use for this population (López, 2022; Yarin, 2022)

Table 4

Implementer Characteristics

Study	Total	Role	Years Experience in Role	Experience with Implemented SEL program	Race/Ethnicity	Gender or Sex Assigned at Birth
Castro-Olivo and Merrell (2012)	N = 2	Teachers	Missing	"Minimal training"	Missing	Missing
Castro-Olivo (2014)	N = 2	Teachers	Missing	4-hour training	Latina (100%)	Women (100%)
Castro-Olivo et al. (2022)	N = 2	Teachers	Missing	4-hour training	Latina (100%)	Women (100%)
Cramer and Castro-Olivo (2015)	N = 2	Graduate students	Missing	4-hour training	Missing	Missing
Elek et al. (2010)	Missing	Missing	Missing	Missing	Missing	Missing
Graves et al. (2017)	N = 3	Doctoral graduate students (n = 2) Professor/graduate student (n = 1)	Missing	3-day training	Missing	Missing
Hecht et al. (2018)	N = 32	Teachers	Missing	1 day training	Missing	Missing

Study	Total	Role	Years Experience in Role	Experience with Implemented SEL program	Race/Ethnicity	Gender or Sex Assigned at Birth
Hopson (2008)	N = 6	Teachers or counselors	Missing	60–90-minute training	Missing	Missing
Kurtz (2020)	N = 10	Teachers (n = 7) Researchers (n = 3)	**Teachers:** 1 year (n = 1) 5 years (n = 1) 6 years (n = 2) 7 years (n = 1) 11 years (n = 1) 19 years (n = 1) **Researchers:** Missing	No experience	Black/African American/Hispanic (14%) White (86%)	Missing
Martinez (2011)	N = 8	Clinical psychology doctoral students	Missing	Missing	Missing	Missing
Moran et al. (2019)	Missing	Graduate students	Missing	Missing	Missing	Missing

Note. MISSING: The information is not mentioned in the article at all; NOT EXAMINED: The authors explicitly mention they did not analyze a variable; NOT REPORTED: The authors had a measure, but did not report findings; NOT REPORTED AS EXPECTED: The way the data were reported were not in the same unit of study as other variables (e.g. reports demographic for the entire school instead of for the group)

Table 5

Quality Appraisal

Authors (Year)	A	B	C	D	E	F	G	H	I	J	K	L	M	Total
Castro-Olivo and Merrell (2012)	1	3	3	3	2	3	3	3	1	0	3	3	3	31
Castro-Olivo (2014)	3	3	2	3	2	3	3	2	1	0	3	3	3	31
Castro-Olivo et al. (2022)	3	3	3	3	3	3	3	3	3	3	3	3	2	38
Cramer and Castro-Olivo (2015)	3	3	2	3	2	3	3	3	2	3	3	3	3	36
Elek et al. (2010)	1	2	3	3	1	2	2	1	0	0	3	3	3	24
Graves et al. (2017)	1	3	2	3	3	3	3	2	3	0	3	3	2	31
Hecht et al. (2018)	3	3	2	3	3	1	3	3	3	3	3	3	3	36
Hopson (2008)	3	3	2	3	3	3	3	3	3	3	3	3	3	38
Kurtz (2020)	2	3	3	3	3	3	3	3	3	3	3	3	3	38
Martinez (2011)	3	3	2	3	2	3	3	3	3	0	3	3	3	34
Moran et al. (2019)	1	3	2	3	3	3	3	3	3	3	3	3	3	36

Note. Per item low score = 0; Per item high score = 3; Max score = 39

A = Theoretical or conceptual underpinning to the research; **B** = Statement of research aim(s); **C** = Clear description of research setting and target population; **D** = The study design is appropriate to address the stated research aim(s); **E** = Appropriate sampling to address the research aim(s); **F** = Rationale for choice of data collection tool(s); **G** = The format and content of data collection tool is appropriate to address the stated research aim(s); **H** = Description of data collection procedure; **I** = Recruitment data provided; **J** = Justification for analytic method selected; **K** = The method of analysis was appropriate to answer the research aim(s); **L** = Evidence that those involved or impacted by the research have been considered in research design or conduct; **M** = Strengths and limitations critically discussed; **Total** = Total score on the QUADS

Table 6

SEL Program Descriptives

Authors (Year)	SEL Program/ Adapted Name (if applicable)	Program Type (as used in the study)	Number of Lessons	Length of Program	Targeted CASEL competencies/other competencies
Castro-Olivo and Merrell (2012)	Strong Teens/Jóvenes Fuertes	Class-wide	12 lessons	12 weeks◆	Self-management* (e.g. stress reduction) Self-awareness* (e.g. emotional awareness) Responsible decision-making* (e.g. problem solving) Relationship skills* (e.g. problem solving) Social awareness* (e.g. empathy training)
Castro-Olivo (2014)	Strong Teens/Jóvenes Fuertes	Class-wide	12 lessons	12 weeks	Self-management* (e.g. dealing with anger) Self-awareness* (e.g. clear thinking) Responsible decision-making* (e.g. problem solving) Relationship skills* (e.g. problem solving) Social awareness* (e.g. understanding other's emotions)
Castro-Olivo et al. (2022)	Strong Teens/Jóvenes Fuertes	Class-wide	12 lessons	12 weeks	Self-management Self-awareness Responsible decision-making Relationship skills Social awareness
Cramer and Castro-Olivo (2015)	Strong Teens	Small group	12 lessons	6 weeks	Self-management Self-awareness Responsible decision-making Relationship skills

Authors (Year)	SEL Program/ Adapted Name (if applicable)	Program Type (as used in the study)	Number of Lessons	Length of Program	Targeted CASEL competencies/other competencies
					Social awareness
Elek et al. (2010)	Keepin' it REAL/keepin' it REAL-acculturation enhanced	School-wide	12 lessons	Not specified (45 mins/lesson)	Self-management* (e.g. refusal self-efficacy) Self-awareness* (e.g. learning normative beliefs) Responsible decision-making* (e.g. model resistance strategies) Social awareness* (e.g. learning normative beliefs) /anti-drug expectancies
Graves et al. (2017)	Strong Start	Small group	14 lessons	14 weeks	Self-management* (e.g. handling uncomfortable emotions) Self-awareness* (e.g. identify body, facial, and situational cues) Responsible decision-making* (e.g. solving problems with friends) Relationship skills* (e.g. making and keeping friends) Social awareness* (e.g. labeling others emotions)
Hecht et al. (2018)	Keepin' it REAL/keepin It REAL – Rural Adaptation	Class-wide	Missing	Missing	Responsible decision-making* (e.g. preventing substance use)

Authors (Year)	SEL Program/ Adapted Name (if applicable)	Program Type (as used in the study)	Number of Lessons	Length of Program	Targeted CASEL competencies/other competencies
Hopson (2008)	Keepin' it REAL	Class-wide	6 lessons	6 weeks	Self-management* (e.g. refusal strategies) Self-awareness* (e.g. explain reasons for refusing) Responsible decision-making* (e.g. avoid substance situations) Relationship skills* (e.g. explain reasons for refusing) Social awareness* (e.g. avoiding substance use situations)
Kurtz (2020)	Strong Start and Strong Kids	Class-wide and small group	12 lessons	12 weeks	Self-management* (e.g. coping with stress) Self-awareness* (e.g. understanding personal emotions) Responsible decision-making* (e.g. clear thinking) Relationship skills* (e.g. problem solving) Social awareness* (e.g. understanding others emotions)
Martinez (2011)	Spark Program	Small group	Not specified, weekly, 45-60 minute lessons	1 school year	Self-management* (e.g. setting high, reasonable goals) Self-awareness* (e.g. anger management) Responsible decision-making* (e.g. resisting peer pressure) Relationship skills* (e.g. building trust) Social awareness* (e.g. using communication skills)

Authors (Year)	SEL Program/ Adapted Name (if applicable)	Program Type (as used in the study)	Number of Lessons	Length of Program	Targeted CASEL competencies/other competencies
Moran et al. (2019)	STAC (Stealing the show, Turning it over, Accompanying others, and Coaching compassion)	Small group	3 lessons	2 hours	Self-management* (e.g. using humor to interrupt bullying) Self-awareness* (e.g. understanding bystander roles) Responsible decision-making* (e.g. getting help from adult) Relationship skills* (e.g. offering support to target of bullying) Social awareness* (e.g. knowing why students bully)

Note. ◆ indicates information received by emailing the first author. * Indicates implicit discussion of competency. Some examples may be repeated as several CASEL competencies were suitable for the skills these programs taught

Table 7

Study Treatment Integrity

Authors (Year)	Treatment Integrity	
	Measure	Outcome
Castro-Olivo and Merrell (2012)	12-item observation tool completed by the first-author over 25% of lessons Audiotapes of sessions—teachers determined amount of time spent on each concept Teacher survey	100% (related to the 12-item observation) Teachers spent recommended amount of time for all main concepts; measured by (related to the audiotapes of sessions) Teachers felt they were "good" to "very good" at managing time to achieve lesson agendas and skill development, and teachers felt their overall effectiveness at teaching the program was "good to very good" (related to the teacher survey)
Castro-Olivo (2014)	Checklist of all main goals for each lesson, completed by principal investigator	Not reported
Castro-Olivo et al. (2022)	Authors state the original study (see: Castro-Olivo (2014)) measured fidelity for each lesson.	Not reported
Cramer and Castro-Olivo (2015)	Missing	Missing

Authors (Year)	Treatment Integrity	
	Measure	Outcome
Elek et al. (2010)	Observers (no further detail was provided)	Qualitative report only: In general, observers rated implementation quality uniformly high. Teachers did not provide consistent student attendance reports.
Graves et al. (2017)	Unspecified means of measurement: First author completed fidelity checks for approximately 20% of all randomly selected sessions	100%
Hecht et al. (2018)	Missing	Missing
Hopson (2008)	The researcher conducted several, brief, 15-minute observations, and implementer interviews were conducted in person or via email	Qualitative report only: In general, most schools implemented with fidelity. One school did not do activities as student interaction was not allowed; two schools changed a few scenarios to make them more relatable; one school did not finish implementing the curriculum
Kurtz (2020)	Measured by number of partially and fully implemented steps divided by total number of applicable intervention steps for each session and multiplied by 100 for each group/teacher	Overall fidelity (M = 97.14%, SD = 2.59)
Martinez (2011)	Not reported	Not reported
Moran et al. (2019)	Missing	Missing

Note. MISSING: The information is not mentioned in the article at all; NOT EXAMINED: The authors explicitly mention they did not analyze a variable; NOT REPORTED: The authors had a measure, but did not report findings;

NOT REPORTED AS EXPECTED: The way the data were reported were not in the same unit of study as other variables (e.g. reports demographic for the entire school instead of for the group)

Table 8

Adaptations and Adaptation Purposes

Authors (Year)	Adaptation Type(s)	Adaptation Purpose(s)
Castro-Olivo and Merrell (2012)	**Surface: Bilingual implementer** **Surface: Materials translated to Spanish** **Surface: Bicultural implementer** **Deep: Incorporating cultural values in SEL program design** -All examples adapted to reflect cultural values, costumes, and traditions **Deep: Incorporating cultural values in SEL program implementation** -Teaching program's target skills altered to reflect cultural reframing often used by Latine populations (e.g. respeto and familismo) **Deep: Other** -Included metaphors and folktales -Taught concepts of acculturative stress and ethnic pride (brings in cultural reality)	Created for a more appropriate use with a Latine Immigrant population
Castro-Olivo (2014)	**Surface: Bilingual implementer** **Surface: Intervention and materials translated into Spanish** **Surface: Bicultural implementer** **Deep: Incorporating cultural values in SEL program design** -All examples were adapted to reflect cultural values, costumes, and traditions **Deep: Incorporating cultural values in SEL program implementation**	Wanted to validate the Jóvenes Fuertes with a Latine ELL sample enrolled in middle and high school students

Authors (Year)	Adaptation Type(s)	Adaptation Purpose(s)
	-Teaching program's target skills altered to reflect the cultural reframing often used by Latine populations (e.g. respeto and familismo **Deep: Other** -Included metaphors and folktales -Taught concepts of acculturative stress and ethnic pride (brings in cultural reality)	
Castro-Olivo et al. (2022)	**Surface: Bilingual implementer** **Surface: Intervention and materials translated into Spanish** **Surface: Bicultural implementer** **Surface:** **Deep: Incorporating cultural values in SEL program design** -All examples adapted to reflect cultural values, costumes, and traditions **Deep: Incorporating cultural values in SEL program implementation** -Teaching program's target skills altered to reflect the cultural reframing often used by Latine populations (e.g. respeto and familismo) **Deep: Other** -Included metaphors and folktales -Taught concepts of acculturative stress and ethnic pride (brings in cultural reality)	Aim was to validate the Jóvenes Fuertes with a Latino ELL sample enrolled in middle and high school students
Cramer and Castro-Olivo (2015)	**Surface: Other** -Used culturally appropriate language (choice between English and Spanish)	Conducted the cultural adaptations for a predominantly diverse group of culturally and linguistically diverse (CLD) students

Authors (Year)	Adaptation Type(s)	Adaptation Purpose(s)
	-Implementers trained in identifying and being sensitive to needs of the group **Deep: Incorporating cultural values in SEL program design** -Used cultural knowledge to align intervention procedures to increase acceptability **Deep: Incorporating cultural values in SEL program implementation** -Encouraged students to consider their own cultural values, customs, and traditions -Considered cultural values that relate to goals for home and school -Considered culture-specific life and family circumstances (e.g. immigration status) **Deep: Other** -Used cultural metaphors of the target group and explain use if not understood by other groups -Taught concepts of acculturative stress and ethnic pride (brings in cultural reality)	who were no longer identified/classified as English language learners (ELL's)
Elek et al. (2010)	**Deep: Incorporating cultural values in SEL program design** -Promoted relevant protective cultural values (such as family centeredness or familismo to reflect the study schools predominantly Mexican and Mexican American populations) **Deep: Other** -Added two additional lessons centering acculturation -Emphasized cultural diversity and ethnic identity as strengths	Addressed the specific needs of students experiences acculturation

Authors (Year)	Adaptation Type(s)	Adaptation Purpose(s)
	-Examined the impact of language on drug resistance and discussions with parents -Explored the changes in identity and values that may occur through acculturation	
Graves et al. (2017)	**Surface: Other** -Implemented at the Tier II instead of Tier 1 level **Deep: Incorporating cultural values in SEL program implementation** -Made literature modifications	Enhanced the curriculum relevance for African American sample Skills taught in the intervention aligned well with teacher desired skills for students to learn
Hecht et al. (2018)	**Deep: Incorporating cultural values in SEL program design** -Collaborated with rural teachers to integrate rural culture surrounding drug use into the curriculum (including role play activities, decision scenarios, and homework, curriculum design, and intervention strategy)	Aim was to integrate information into the program on how rural cultures differ from urban and suburban counterparts (ie. health disparities, weaker protective factors)
Hopson (2008)	**Surface: Other** -Adapted teacher version of the manual to make 6 sessions **Deep: Incorporating cultural values in SEL program design** -Held focus groups and group activities to describe personal stories -Student lead scenario development -Recreated videos that fit student experiences (recreated by students) -Created new role play scenarios (led by students)	Collapsed the curriculum into a six-week session as alternative school sessions typically only go for 6-8 weeks, so a 10-week intervention was not feasible or culturally appropriate Created an adapted program to be culturally grounded for youth in Texas Made the curriculum appropriate for older age students (14-19 year olds) in alternative schools

Authors (Year)	Adaptation Type(s)	Adaptation Purpose(s)
Kurtz (2020)	**Surface: Translation of materials** -Provided the book "The Way I Feel" in Spanish -Family bulletin translated into Spanish to send home with students **Surface: Inclusion of ethnic lifestyle elements** -Asked students to share synonyms for angry in the language they use at home **Surface: Inclusion of same race/ethnicity role model** -Chose an alternative book where the main character is a child of color: Marvelous Me by Lisa Bullard -Chose an alternative book the main character is a child of color: Those Shoes by Maribeth Boelts **Surface: Other** -Used individual body maps to identify how feelings are experienced in the body -Included additional mindfulness activity and coping skill instruction (e.g., breathing, squeezing clay) **Deep: Incorporating cultural values in SEL program design** -Got community feedback on researcher created adaptations -Included culturally relevant examples for having multiple feelings at once (e.g., when someone has to leave Boston) -Included psychoeducation about how experiencing trauma can make it difficult to identify feelings when all feelings can feel like angry or scared **Deep: adjusting materials to literacy level of participant**	Focus group was implemented to guide the intervention adaptation process Focused on internalizing problems in the context of a high needs urban elementary school

Authors (Year)	Adaptation Type(s)	Adaptation Purpose(s)
	-Incorporated extension activity "Mirror Mirror" to support English Language Learners who may require more visuals **Deep: Other** -Asked students to share physical clues related to how people at home show different emotions -Taught the trauma response of "flight, fight, or freeze" -Discussed the link between trauma and anger, and how trauma can make us feel like many situations/experiences are "big deals" despite them being "little deals" -Included psychoeducation about how experiencing trauma can make it difficult to identify physical clues of emotions in others	
Martinez (2011)	**Surface: other** -Lengthened course of treatment (to support community with high need) -Multiple modes of treatment options (started group, gave option for individual support) **Deep: Involving family in SEL program** -Formal graduation party for students and their families **Deep: Other** -Advocacy efforts (e.g. involved in IEP meetings, translation services)	Adapted to fit the specific needs of the diverse, at-risk populations of children and youth who attended two local public schools (one elementary school and one junior high school) in the Bay Area Extended treatment: schools had limited resources and high need
Moran et al. (2019)	**Surface: Bilingual implementer** -Recruited ethnically diverse and bilingual graduate students who spoke English and Spanish **Surface: Inclusion of ethnic lifestyle elements** -Infusion of culturally relevant language (e.g. used "defender" instead of "advocate")	In response to the growing need for relevant and accessible interventions in ethnically blended, rural, low-income schools, the researchers adapted the STAC program to be culturally appropriate for this population

Authors (Year)	Adaptation Type(s)	Adaptation Purpose(s)
	Surface: Other -Trained delivery consistent with cultural norms and values -Trainers encouraged to be intentional about fostering connections with and between students to promote relationship building **Deep: Incorporating cultural values in SEL program design** -Inclusion of culturally relevant bullying experiences and training goals -Included examples related to culture and ethnicity	

Note. The adaptation purposes were extracted to retain as much of the studies' original verbiage as possible. Bolded headings are the common features of surface and deep level features. Listed below are extracted examples from the reviewed studies.

Table 9

Student Outcomes

Authors (Year)	Dependent Variables	Measures	Outcomes of Measures (Quantitative, if possible, otherwise qualitative provided)	Author Description of Findings
Castro-Olivo and Merrell (2012)	-Knowledge of the SEL content taught in program -Participant internalizing symptoms -Acculturative stress -Students' sense of belonging/connectedness to school	- Internalizing symptoms test -Strong Teens knowledge test -Societal attitude familial environments-for children (SAFE-C; Spanish version) -People in my life scale (sense of school belonging subscale)	**-Knowledge of SEL** content taught in the program: (*d* = 0.98, large effect)*** **-Internalizing disorders:** (*d* = 0.15) -**Acculturative stress:** (*d* = 0.20, small effect) -**School belonging:** (*d* = 0.37, small effect)	-**Knowledge of SEL:** Statistically significant SEL knowledge was gained from baseline to post-intervention. -**Internalizing disorders:** No statistical significance was found for students' reported levels of symptoms of internalizing disorders. -**Acculturative stress:** No statistical significance was found for students' reported levels of symptoms of internalizing disorders. -**School belonging:** No statistical significance was found for students' reported levels of symptoms of sense of school belonging.
Castro-Olivo (2014)	-Social emotional resiliency -Student knowledge of SEL concepts	-BERS-2 -Strong Teens Knowledge Test—Spanish version	-**Social emotional resiliency:** (*d* = .44)* -**Student knowledge of SEL concepts** (*d* = 1.27)*	Students assigned to the intervention condition demonstrated significantly more resilience and knowledge.

Authors (Year)	Dependent Variables	Measures	Outcomes of Measures (Quantitative, if possible, otherwise qualitative provided)	Author Description of Findings
Castro-Olivo et al. (2022)	-Self-awareness -Relationship skills -Self-management -Social awareness -Responsible decision-making	- BERS-2 (Behavioral and emotional rating scale) Spanish version, restructured for this study	Intervention vs control group: (*d* = pre intervention/post intervention) -▲ Self-awareness (*d* = .23/.662)* -▲Relationship skills (*d* = -0.0017/0.39)** -▲Self-management (*d* = 0.187/0.0725) -▲Social awareness (*d* = 0.1799/0.2276) -▲Responsible decision-making (*d* = 0.0638/.1043)	Students in the intervention condition showed significantly higher-levels of self-awareness and relationship skills.
Cramer and Castro-Olivo (2015)	-Student resiliency -Internalizing symptoms	- BERS-2 YRS (Behavioral and emotional rating scale, youth rating scale) -CBCL, youth self-report only (24 items only, not full scale)	-**Student resiliency:** Before and after intervention scores: t(19) = −2.12, p = 0.048, ES = 0.2); post-intervention and follow-up scores: t(19) = −0.23, p = 0.822 -**Internalizing symptoms:** Before and after intervention scores: t(19) = 0.82, p = 0.424); post-intervention and follow up	-**Student resiliency:** There was a significant difference between resiliency scores before and after the intervention. There was no significance between post-intervention resiliency scores and follow-up resiliency scores was not significant. -**Internalizing symptoms:** There was not a statistically significant difference between before and after intervention scores. There was also not a

Authors (Year)	Dependent Variables	Measures	Outcomes of Measures (Quantitative, if possible, otherwise qualitative provided)	Author Description of Findings
			scores: t(19) = 1.17, p = 0.256	statistically significant difference between post-intervention and follow-up internalizing symptom scores.
Elek et al. (2010)	- Lifetime substance use prevalence -Past month's prevalence -Intentions to use substances -Refusal efficacy -Hypothetical alcohol resistance -Number of substance use resistance strategies -Personal anti-drug norms (ie. rejection of peer substance use) -Descriptive substance use norms (ie. perception of prevalence of	- Long scannable questionnaire (104-177 items) with items adapted from Kasen, Vaughan, and Walter (1992), Hecht et al. (2003), and Hansen and Graham (1991)	Effect sizes (5th grade only vs control/ 7th grade only vs control/ 5th+7th vs control) -**Past month's substance use prevalence** (*d* = .48*/-.23/-0.05) - **Lifetime substance use prevalence** (*d* = .11/-.20/-0.18) -**Intentions to use substances** (*d* = .02/-.09/-0.15) **-Refusal efficacy** (*d* = -0.15/-0.08/0.15) -**Hypothetical alcohol resistance** (*d* = 0.06/0.16/0.16) -**Number of strategies** (*d* = 0.13/0.04/0.15) -**Personal anti-drug norm** (*d* = -0.06/-0.11*/-0.06)	**-Past month substance use:** The increase from baseline to wave 6 (ie. 8th grade) in the prevalence of **past month's use** reported by control condition students (from 43% to 56%) was significantly less than that reported by students receiving the kiR-AE (keepin' it real-acculturation enhanced) curriculum in only the 5th grade (from 29% to 78%). -**Lifetime substance use:** The increase from baseline to wave 6 (ie. 8th grade) reported by control students did not differ significantly from that reported by intervention students, except in the case of students receiving the kiR-AE version in only the 5th grade. Students in that condition increased in their **lifetime prevalence of substance** use (from 23% to

Authors (Year)	Dependent Variables	Measures	Outcomes of Measures (Quantitative, if possible, otherwise qualitative provided)	Author Description of Findings
	peer substance use) -Positive substance use expectancies		-**Descriptive substance use norms** (ie. views that peers are using substances) (*d* = -0.02/-0.26/-0.21) **-Positive substance use expectancies** (*d* = 0.07/0.11/-0.34)	71%) significantly more than did control condition students (from 30% to 63%) Students receiving the curriculum versions in the 7th grade or in both the 5th and 7th grades did not appear different from the control group with regards to recent substance use.
Graves et al. (2017)	-Social-emotional knowledge -Social-emotional assets -Externalizing behaviors	-Strong Start Content Knowledge Assessment (SSCK) -Social-Emotional Assets and Resiliency Scales (SEARS) -BASC-2	-**Social-emotional knowledge:** 16% increase, *d* not calculable -**Social-emotional assets:** self-regulation: *d* = 0.99 large self-competence: *d* = 1.38 large empathy: *d* = 0.47 small responsibility: *d* = 0.10 small) -**Externalizing behaviors** (*d* = 0.36 small)	-**Social-emotional knowledge:** "From pretest to posttest, the students' scores increased on the Strong Start Content Knowledge assessment. Between pre- and posttest, students demonstrated an increase in their social-emotional knowledge. Specifically the group average was 68% correct on the pretest and 84% on the posttest." -**Social-emotional assets:** **"**Students' self regulation and self-competency improved over the course of the intervention." Empathy and responsibility did not have statistically significant

Authors (Year)	Dependent Variables	Measures	Outcomes of Measures (Quantitative, if possible, otherwise qualitative provided)	Author Description of Findings
				differences between before and after scores of the intervention. -**Externalizing behaviors:** "There were not any statistically significant decreases in teacher rated externalizing problems…the change in student's aggression as measured by effect sizes was not practically meaningful."
Hecht et al. (2018)	-Youth lifetime alcohol use -Youth lifetime cigarette use -Youth lifetime marijuana use -Youth lifetime chewing tobacco use	-4 items adapted from Hansen and Graham, 1991)	▲**Wave 1-Baseline:** (*d* = high quality adapted vs control/low quality adapted vs control) -Alcohol use (*d* = 0.07/0.145) -Cigarette use (d = 0.08/0.16) -Marijuana use (*d* = 0.068/0.127) -Chewing tobacco use (*d* = -0.104/-.117) ▲**Wave 4- 24 months after implementation):** (*d* =high quality adapted vs control/low quality adapted vs control) -Alcohol use (*d* = -.14/-.11)	-Youth in both standard and rural adaptation programs reported significantly less **cigarette use** than those in the control condition. -Programs with high quality delivery in the multicultural/rural conditions were in the desired direction but not statistically significant. -Youth in both high- and low-quality delivery of the adapted/rural condition reported significantly less **cigarette use** than those in the control condition. -All the other effects for high-quality adapted/rural

Authors (Year)	Dependent Variables	Measures	Outcomes of Measures (Quantitative, if possible, otherwise qualitative provided)	Author Description of Findings
			-Cigarette use (*d* = -.20/-.13)* -Marijuana use (*d* = -0.007/0.0569) -Chewing tobacco (*d* = -.169/-.20)	curriculum were in the desired direction but not statistically significant. -Effects for low-quality adapted/rural curriculum were not statistically significant but were in the desired direction for **alcohol use** and **chewing tobacco use** -Youth in the high-quality, nonadapted/urban condition reported less **marijuana use** than those in the low quality, non-adapted/urban condition but neither high nor low quality delivery of the non-adapted curriculum produced substance use outcomes that differed significantly from the control.
Hopson (2008)	-Past month alcohol use -Past month marijuana use -Intentions to accept alcohol, -Intentions to accept marijuana	-Texas School Survey of Substance Use (3 items adapted for past month alcohol use) -Texas School Survey of Substance Use (1 item adapted	-**Alcohol use** by group (η_p^2 = 0.097) -**Alcohol use** by group x time (η_p^2= 0.108)* -**Alcohol use** by group x time x age (η_p^2 = 0.295)** -**Alcohol use** by time (η_p^2= 0.164) **	-There was a significant difference in **alcohol use** between groups over time (F (2, 60) = 3.637, p<.05). -**Alcohol use** accounted for 10.8% of the difference in group means over time. -**Alcohol use:** The age difference between groups may

Authors (Year)	Dependent Variables	Measures	Outcomes of Measures (Quantitative, if possible, otherwise qualitative provided)	Author Description of Findings
		for past month marijuana use) -Drug Resistance Strategies (DRS) Project (2 items adapted for intentions to accept alcohol) -Drug Resistance Strategies (DRS) Project (1 item adapted for intentions to accept marijuana)	-**Marijuana use** by group ($\eta_p^2 = 0.007$) -**Marijuana use** by group x time ($\eta_p^2 = 0.042$) -**Marijuana use** by group x time x age ($\eta_p^2 = 0.057$) -**Marijuana use** by time (partial eta squared = 0.044) -**Intentions to accept alcohol** by group ($\eta_p^2 = 0.24$) -**Intentions to accept alcohol** by group x time ($\eta_p^2 = 0.23$) *** **-Intentions to accept by alcohol** group x time x age ($\eta_p^2 = 0.142$) **-Intentions to accept alcohol** by time ($\eta_p^2 = 0.037$) -**Intentions to accept marijuana** by group ($\eta_p^2 = 0.001$) **-Intentions to accept marijuana** by group x time ($\eta_p^2 = 0.034$)	have influenced the difference in mean scores over time. -**Alcohol use** decreased over time for the experimental group and remained fairly stable for the comparison group, indicating that the experimental group reported significantly greater decreases in alcohol use than the comparison group -The groups did not differ in their **marijuana use** over time -Although the difference was not significant, the experimental group's **marijuana use** decreased steadily over time, while the comparison group's use decreased slightly and then increased **-Intentions to accept alcohol:** Experimental group was more likely to disagree that they would accept an offer for alcohol than the comparison group over time **-Intentions to accept marijuana:** The experimental

Authors (Year)	Dependent Variables	Measures	Outcomes of Measures (Quantitative, if possible, otherwise qualitative provided)	Author Description of Findings
			-Intentions to accept marijuana by group x time x age (η_p^2= 0.059) **-Intentions to accept marijuana** by time (η_p^2= 0.004)	group scores increased over time while the scores for the comparison group decreased. Since the analysis suffers from lack of power due to small sample and violation of assumptions, the trend warrants further study
Kurtz (2020)	-Students' internalizing behaviors	- The Behavior Intervention Monitoring Assessment System 2 (BIMAS-2)	**Negative affect scores:** from pre intervention to post intervention -Class wide condition: *d* = -1.23 -Small group condition: *d* = -2.87 -Receiving special education: *d* = -2.87 -General education: *d* = -1.14	-"Decreases in Negative Affect T scores suggest improvement in symptoms/functioning while increases in scores may reflect worsening symptoms (McDougal et al., 2011). Negative Affect T scores decreased from pre-intervention to post-intervention for both the classwide, pre-intervention: *M* = 65.42 , *SD* = 5.71 post-intervention: *M* = 58.42, *SD* = 7.35, and small group conditions, pre-intervention: *M* = 66.68, *SD* = 5.49, post-intervention: *M* = 56.63, *SD* = 9.49, suggesting teachers rated their students as having fewer problem behaviors related to

Authors (Year)	Dependent Variables	Measures	Outcomes of Measures (Quantitative, if possible, otherwise qualitative provided)	Author Description of Findings
				Negative Affect after the intervention."
Martinez (2011)	-Educational aspirations -Adjustment (psychological) -Anger management -Self-esteem -Ethnic pride -Self-reported peer attachments -Gang involvement	- Junior High life Experiences Survey (JHLES) -Survey of Children's Exposure to Community Violence (SECV-SR) -Educational Aspirations (researcher designed) -Trauma Symptom Checklist for children (TSCC) -Anger management (researcher designed + vignettes from Slaby & Wildon-Brewer 1992) -Rosenberge Self-Esteem Scale -Multigroup Ethnic Identity Measure (MEIM)	**-Educational aspirations:** [F(1,29) = 1.42, p = .24] -**Psychological adjustment:** [F(1,29) = .01, p = .94] -**Anger management:** "From pre-group to post-group, participants said they would be less likely to walk away and more likely to try to find out why people were fighting." -**Self-esteem:** [F(l,29) = 2.11, p = .16] -**Ethnic pride:** [F(l,29) = 2.11, p = .16] -**Peer attachment:** Elementary students: [F(l, 16) = 6.06, p = .026]; Junior high school students: [F(l, 13) = .49, p = .50]") -**Gang involvement:** [F(l,29) = 1.55, p = .22]	**-Educational aspirations:** no significant pre-group to post-group change -**Psychological adjustment:** no significant pre-post group change -**Anger management:** provided in Outcomes of Measures column **-Self-esteem:** "No significant changes from pre- to post-group in self-esteem…Participants tended to report they agree to feeling self-esteem." -**Ethnic pride:** There were no significant changes from pre- to post-group in ethnic pride. Participants tended to report that they agree to feeling ethnic pride -**Peer attachment:** "Participants in elementary school showed an improvement in their reported peer attachment from pre to post-

Authors (Year)	Dependent Variables	Measures	Outcomes of Measures (Quantitative, if possible, otherwise qualitative provided)	Author Description of Findings
		-Peer Scale of the Inventory of Parent and Peer Attachment (IPPA) -Gang Involvement (researcher designed)		group…However, there was no significant change for junior high school participants" -**Gang involvement:** "Participants tended to report that they did not find it that necessary to belong to a gang. There were no significant changes from pre- to post-group in gang involvement"
Moran et al. (2019)	-Knowledge and confidence to intervene -Using "stealing the show" strategy -Using "turning it over" strategy -Using "accompanying others" strategy -Using "coaching compassion" strategy	-Student-Advocates pre- and post-scale -Use of STAC strategies (each strategy measured by 1, 5 point-Likert item)	Results from full sample: -**Knowledge and confidence to intervene** by time (η_p^2 = .68) *** -**Knowledge and confidence to intervene** by time x ethnicity (η_p^2 = .02) **-Using "stealing the show" strategy** (white vs Hispanic student use: φ = .39)* -**Using "turning it over" strategy** (white vs Hispanic student use: φ = .02) -**Using "accompanying others" strategy** (white vs	- Results indicated Hispanic students used the strategy "stealing the show" more frequently than White students, with no significant differences between White and Hispanic students for the other three strategies

Authors (Year)	Dependent Variables	Measures	Outcomes of Measures (Quantitative, if possible, otherwise qualitative provided)	Author Description of Findings
			Hispanic student use: φ = .02) -**Using "coaching compassion" strategy** (white vs Hispanic student use: φ = .02)	

Note. Where possible, author language was used. Effect sizes were extracted from papers. ▲ symbolizes calculations made by the authors of this systematic review. * indicates p < .05; ** indicates p< .01; *** indicates p < .001 φ indicates partial eta squared, which ""means the proportion of the variance that is attributable to the effect of the dependent variable" (Hopson, 2008, p. 136) *d* = indicates the Cohen's D value, which demonstrates the number of standard deviations between the two means. In Cramer and Castro-Olivio (2015), the effect size type calculated by the authors was not reported, thus the authors' language of "ES" was used in this review.

Table 10

Social Validity and Outcomes

Authors (Year)	Social Validity Construct(s)	Measure	Outcome
Castro-Olivo and Merrell (2012)	-Social validity (student): acceptability and satisfaction -Social validity (teacher): perceived acceptability, satisfaction, feasibility	-Social validity: student report (researcher developed, 7-item scale) -Social validity: teacher report (researcher developed, 16-item scale)	-**Student acceptability: Not Reported** -**Student satisfaction:** highly satisfied (80%+) -**Teacher perceived acceptability:** both said students liked the intervention -**Teacher satisfaction:** very satisfied -**Teacher feasibility:** "good" to "very good"
Castro-Olivo (2014)	-Acceptability -Satisfaction	Social validity: student report (researcher developed, 9-item scale)	-**Acceptability:** (high, 84.4%+ endorsed statements) -**Satisfaction:** Not Reported
Castro-Olivo et al. (2022)	Not reported	BERS-2 (restructured) (student completed)	Not reported
Cramer and Castro-Olivo (2015)	-Acceptability -Satisfaction -Overall social validity	9-item measure based on Wolf's (1978) definition of social validity. Scale was validated in previous studies examining a culturally adapted version of Strong Teens (student only)	-**Acceptability:** Not reported as expected -**Satisfaction:** Not reported as expected -**Overall social validity:** Not reported as expected -In general, high ratings on all 9 items. Ranged from 80-100% agreement (SD = 0.07).
Elek et al. (2010)	MISSING	MISSING	MISSING

Authors (Year)	Social Validity Construct(s)	Measure	Outcome
Graves et al. (2017)	- Social significance, social importance, social appropriateness -Teacher view of social validity	-Student social validity questionnaire -Teacher survey -Teacher interview	-**Social significance, social importance, social appropriateness**: high endorsements of all items -**Teacher view of social validity**: viewed the intervention as a good first step, but needed a stronger focus on Black boys, resilience, and needed more time
Hecht et al. (2018)	MISSING	MISSING	MISSING
Hopson (2008)	MISSING	MISSING	MISSING
Kurtz (2020)	-Teachers' acceptability and likelihood of continued implementation of the adapted intervention	-The Usage Rating Profile Intervention-Revised (URP-IR), teacher rated	Max score = 6 -**Acceptability** (*M* = 5.32, *SD* = 0.96) -**Understanding** (*M* = 3.92, *SD* = 0.95) -**Home school collaboration** (*M* = 4.86, *SD* = 0.91) -**Feasibility** (*M* = 4.78, *SD* = 0.79) -**System climate (***M* = 5.50, *SD* = 0.56) -**System support** (*M* = 4.19, *SD* = 1.25)
Martinez (2011)	-Overall perceived effect -Overall program evaluation	-Perceived effect scale (student rated) Likert Scale: 1 = Not at all; 5 = Very helpful -Post interview (students only)	-**Overall perceived effect**: *N* = 31, *M* = 3.56 (.84) -**Overall program evaluation** (N = 31, M = 4.54 (.69) -**Group interview:** many of the participants felt an overall appreciation for the program and how it helped them; help with anger management was listed

Authors (Year)	Social Validity Construct(s)	Measure	Outcome
			as the area in which they felt most helped. -Participants reported that they learned most about not bullying others, anger management, educational aspirations, expressing their feelings and one student mentioned culture and racism.
Moran et al. (2019)	Students' experiences participating in the culturally adapted STAC program and using the STAC strategies to intervene in bully situation in a predominantly white and Hispanic middle school located in a low-income rural community	- Participants' experience (student interview)	Results only from qualitative sample: - 47.8% of students spoke about an increase in awareness about bullying and reported using all four strategies to intervene across multiple contexts -39% of students intervened using the STAC strategies despite their fears -56.5% of students reported feeling good about themselves, confident, and feeling like they could better recognize bullying to make a difference at their school after using STAC strategies -43.4% of students reported that since participating in the training and using STAC strategies they are making new friends, building stronger relationships, and becoming more discerning in friendships

Note. "Acceptability" and "satisfaction" may have been the general ways authors were defining social validity, however, for the most clearway of coding the two, the two were separated. MISSING: The information is not mentioned in the article at all; NOT EXAMINED: The authors explicitly mention they did not analyze a variable; NOT REPORTED: The authors had a measure, but did not report findings; NOT REPORTED AS EXPECTED: The way the data were reported were not in the same unit of study as other variables (e.g. reports demographic for the entire school instead of for the group)

Table 11

Numerical Codes for 2nd, 3rd, and 4th Grade Academic Letter Grades

2nd Grade Letter Grades	Study Code	3rd and 4th Grade Letter Grades	Study Code
E+	8	A+	13
E (Excellent)	7	A	12
E-	6	A-	11
S+	5	B+	10
S (Satisfactory)	4	B	9
S-	3	B-	8
I (Improvement Needed)	2	C+	7
X (Area of Concern)	1	C	6
NG (No Grade Given)	777	C-	5
P (Pass)	888	D+	4
NP (No Pass)	999	D	3
		D-	2
		F	1

Table 12

Research Question 4: Joint Display of Teacher Practicality Perceptions for Triangulation

Quantitative		Qualitative
URP-IR Feasibility Subscale Items	Mean (*M*), Standard Deviation (*SD*)	Exemplar Quotes Related to Identified Themes that Either Support or Contradict the Rated Items Below
		Theme 1: BGP Is Convenient for Teachers
3. I would be able to allocate my time to implement this intervention.	*M* = 5.28, *SD* = .95 On average, teachers **agreed to strongly agreed** with this statement.	**Supporting Quote:** "It's consistent. The kids know what to expect so the Monday we [are] doing the lesson and they know what we're going to do and they know the consistency of it. So, kind of like that routine and structure… it is really nice to have and for teachers just simple to implement" (Christina**).
8. The total time required to implement the intervention procedures would be manageable.	*M* = 5.29, *SD* = 1.11 On average, **teachers agreed to strongly agreed** with this statement.	**Contradicting Quote:** "I tried to do that stuff [add more engaging activities] to make it more hands on for the kids. Not sure if I have time because there's there's so much that we have to cover…It's kind of "you have to." Like you just designate that time. Here's this window of time I need to do it" (Susan**).
13. Preparation of materials needed for this intervention would be minimal.	*M* = 5.86, *SD* = .38 On average, **teachers agreed to strongly agreed** with this statement.	**Supporting Quote:** The accessibility, the scope and sequence is good as far as it's something concrete that I can follow. And I can follow it at the drop of a hat. You know, I don't need to do any prep work actually at all other than like, logging in and having it ready. I love that part of it" (Julia**).

17. Material resources needed for this intervention are reasonable.	*M* = 5.57, *SD* = .53 On average, **teachers agreed to strongly agreed** with this statement.	**Contradicting Quote:** "The website itself is like a plethora of information. Although I don't think we use it to its ability because they're just, you know, there isn't like, paid time in the day for teachers to go through it" (Julia**).
19.* This intervention is too complex to carry out accurately	*M* = 3.71, *SD* = 1.60 On average, teachers **slightly disagreed to slightly agreed** with this statement.	**Contradicting Quote:** "Everything is prepared, ready to go…It's linked in And we just click, you know, click on our lesson plans and it's there" (Susan).
27. The amount of time required for record keeping would be reasonable.	*M* = 5.00, *SD* = .82 On average, teachers **agreed** with this statement.	Not Available
Total Feasibility Subscale Score	*M* = 30.71, *SD* = 3.59	------

Note. The Usage Rating Profile-Intervention, Revised (URP-IR) items were rated on a 6-point Likert-type scale where 1 = strongly disagree, 2 = disagree, 3 = slightly disagree, 4 = slightly agree, 5 = agree, and 6 = strongly agree. The highest possible summed rating a teacher could provide on the feasibility subscale was a 36. The asterisk (*) indicates the item was reverse scored. The double asterisk (**) indicates the direct quote has been edited for clarity by removing filler words, linguistic utterances, and/or repeated words; however, the integrity of the quote was not changed. Quotes were considered "Not Available" if there were not quotations that both fit the theme and related to the feasibility item. BGP stands for Be Good People.

Table 13

Research Question 5: Joint Display of Acceptability Perceptions for Triangulation

Quantitative		Qualitative	
URP-IR Acceptability Subscale Items	Mean (*M*), Standard Deviation (*SD*)	Exemplar Quotes Related to Identified Themes that Either Support or Contradict the Rated Items Below	
		Theme 2: BGP is Pertinent for Students	Theme 3: BGP Skill Generalizability is Limited at the Elementary Level
1. This intervention is an effective choice for addressing a variety of problems.	*M* = 4.71, *SD* = .76 On average, teachers **slightly agreed to agreed** with this statement.	**Supporting Quote:** "Consistently, you can just find the resource you need, and teach that specific skill. any point throughout the year" (Laura). **Supporting Quote:** "I think there's a lot of things in there that are good you know, when they're ready for it…I know, we've had some really good discussions with some of the lessons that we've done. So, I think there's a lot of good that comes out of it" (Susan).	**Contradicting Quote:** As far as the students…Do they apply it? Sometimes it seems like they do but sometimes not. But I don't know if there's any curriculum out there that... it's like they can talk the talk in class, but when the scenarios and situations actually happen. It's hard to tell sometimes if they're actually applying those things" (Christina**).
7. The intervention is a fair way to handle the child's behavior problem.	*M* = 4.00, *SD* = 1.53 On average, teachers **slightly agreed** with this statement.	Not Available	Not Available

9. *I would not be interested in implementing this intervention.	*M* = 3.86, *SD* = 1.07 On average, teachers **slightly disagreed to slightly agreed** with this statement.	Not Available	Not Available
11. I would have positive attitudes about implementing this intervention.	*M* = 4.71, *SD* = .49 On average, teachers **slightly agreed to agreed** with this statement.	Not Available	Not Available
12. This intervention is a good way to handle the child's behavior problem	*M* = 3.71, *SD* = 1.38 On average, teachers **slightly disagreed to slightly agreed** with this statement.	Not Available	**Contradicting Quote**: "I think it is not appropriate for first grade. It is too above their heads, like they don't make the connection, doesn't apply to them" (Ava). **Contradicting Quote:** "It is hard to keep kids engaged. When it's like the same thing over and over and over every week" (Natalie).
18. I would implement this intervention with a good deal of enthusiasm.	*M* = 5.14, *SD* = 1.07 On average, **teachers agreed to strongly agreed** with this statement.	Not Available	Not Available

21. This intervention would not be disruptive to other students.	*M* = 4.86, *SD* = 1.46 On average, teachers **slightly agreed to agreed** with this statement.	Not Available	Not Available
22. I would be committed to carrying out this intervention.	*M* = 5.00, *SD* = .82 On average, teachers **agreed** with this statement.	Not Available	**Supporting Quote**: "I think I can teach it with fidelity, but I have to modify…to meet my students' needs" (Ava). **Supporting Quote:** "The younger kids with their being so centered on themselves. I felt like that's where I said I'd have to change some of the scenarios or I would add, you know, ad lib a little bit with the scenarios to try and make it more personal. But, I would say, overall, I mean, I thought it was pretty good. I mean, I try to teach it as it's scripted, you know, try to go through it as much as I can, and cover everything that's there" (Susan).
23. The intervention procedures easily fit in with my current practices.	*M* = 4.57, *SD* = 1.27 On average, teachers **slightly agreed to**	Not Available	Not Available

	agreed with this statement.		
Total Acceptability Subscale	*M* = 40.57, *SD* = 7.09	-----	-----

Note. The Usage Rating Profile-Intervention, Revised (URP-IR) items were rated on a 6-point Likert-type scale where 1 = strongly disagree, 2 = disagree, 3 = slightly disagree, 4 = slightly agree, 5 = agree, and 6 = strongly agree. The highest possible summed rating a teacher could provide on the acceptability subscale was 54. The asterisk (*) indicates the item was reverse scored. The double asterisk (**) indicates the direct quote has been edited for clarity by removing filler words, linguistic utterances, and/or repeated words; however, the integrity of the quote was not changed. Quotes were considered "Not Available" if there were not quotations that both fit the theme and related to the acceptability item.

Table 14

Research Question 6: Relationship Between CRE Curriculum Scorecard and Teacher Interview Quotes Regarding BGP Universality

CRE Curriculum Scorecard Subscale	CRE Curriculum Scorecard Score	Exemplar Quotes (If Available)
Representation Score	**Culturally Insufficient: "**The curriculum likely has culturally and racially ambiguous characters. Few characters and stories are portrayed in a culturally and historically accurate way. There is likely little to no diversity among curriculum contributors and illustrators" (J. Byran-Gooden et al., 2023, p. 23).	"I think it covers pretty much all of our students...they got different pictures, and they've got different, you know, ethnicities, and people are represented in the program" (Christina). "Even like the character name choices, like the characters' names, and the scenarios aren't like 'Mary,' 'John,' like a couple weeks ago, there was 'Juan' and there's other names from other cultures. And that might be one way that like, their culture could be reflected in the Be Good People is just names that they recognize" (Natalie**).
Social Justice Score	**Culturally Insufficient:** "The curriculum predominantly centers white or Eurocentric ideas and culture in most of its components. For the most part students are not encouraged to think critically, or take action to combat inequity. The curriculum provides weak connections to students' lived experiences. There is hardly any opportunity for teachers to engage cultural	"If I had a student that celebrated Kwanzaa in the classroom, which I don't, and I never have, but I think you know, that could be a really good lesson and maybe it could be brought up a couple more times, or even the fact that like, Native American children typically don't make eye contact because it's, it's considered disrespectful and rude. But like, you know, in a

	responsiveness" (J. Byran-Gooden et al., 2023, p. 24).	white family, that's something so simple, but in a first grade, they might not understand that, so you could talk about that it would be like adding those pieces in" (Julia).
Teacher Training Materials Score	**Culturally Insufficient:** "There is a little guidance on engaging diverse learners or culturally responsive teaching in the teacher's materials, but it is mostly on a superficial or symbolic level. It is seen as additive, rather than central to the curriculum and teaching"	Not Available: Teachers primarily discussed the types of CRP training they received. Teachers mostly stated they did not receive any formal training in BGP beyond a meeting with the SEL coordinator and learning how to navigate the BGP website.

Note. A double asterisk (**) indicates the direct quote has been edited for clarity by removing filler words, linguistic utterances, and/or repeated words; however, the integrity of the quote was not changed.

Table 15

Culturally Relevant Pedagogy Training Types Mentioned

Culturally Relevant Pedagogy (CRP) Training Types	Number of Teacher Mentions (*n*), Percentage of Total Teachers (%)	Exemplar Quotes
District-wide training sessions	*n* = 5; 71.4%	"We have what we call [professional development day name]. So, it's a training session that we have at school one day where other schools come. So, this year they came to our school, and we had different presenters for different breakout sessions. And I did take a session on cultural pedagogy" (Susan**).
Informational Speaker	*n* = 4; 57.1%	"He kind of just gave the history of like, on Native American tribes in Minnesota and then talked a little bit about Ojibwe culture, and then language" (Natalie***).
Indigenous student teacher taught about her culture	*n* = 2; 28.6%	"I had a student teacher who's Native American and really, like, I'm not saying that's our only culture, but it's a bigger one in our area is Native American, you know. And so like, my student teacher did a presentation on ribbon dresses and pow wows and artifacts and things like that, to just my class and a preschool class" (Ava**).
Discussed during Master's teacher training program	*n* = 2; 28.6%	"I took a culturally cultural pedagogy masters course" (Natalie).
"Cover Your Bias" Workshop	*n* = 1; 14.3%	"I've been a part of lots of different types [of culturally relevant pedagogy training]. I've been a part of some where it's just like, covering your bias, you know, trying to like get the staff to reflect on their own bias, and then maybe even put into words, what their core beliefs are so that you can see how that might translate to the way you
Sage Ceremony	*n* = 1; 14.3%	

		teach. Workshops like that. I've been a part of workshops where, like, we all sit around and do like a sage ceremony and talk about the way things might go at home for a Native American student" (Julia).
Book study in professional learning communities (PLCs)	*n* = 1; 14.3%	"We're doing different things with PLCs all the time. And sometimes we will have...done book studies…The title was a few years ago, but it was one that was about poverty" (Christina).

Note. A double asterisk (**) indicates the direct quote has been edited for clarity by removing filler words, linguistic utterances, and/or repeated words; however, the integrity of the quote was not changed.

Table 16

Teachers' Cultural Values Mentioned

Teachers' Described Values	Number of Teacher Mentions (*n*); Percentage of Total Teachers (%)	Exemplar Quotes
Respect (e.g., respecting each other's identities, respecting authority)	*n* = 5; 71.4%	"Respect yourself and others, listen to...listen, respect authority" (Laura).
Hard work/Work ethic	*n* = 3; 42.9%	"I think there's this whole idea of like, like hard work gets you places and, you know, you can always rely on that. And whether that's true or not, I think it's like a common belief in most families. You know, so it's really easy to talk to families about that at conferences. They're coming from that background, like, 'well, they just got to work harder' and sometimes you're like, 'they're struggling, for whatever reason, and they do work hard, you know, there is another piece,' but I understand that" (Julia).
Centering family (e.g,. family time, family first, etc.)	*n* = 3; 42.9%	"The importance of family in my life, my time, my family, helping others. Those are all things that I've always been stressed with my children" (Susan)
Religious Faith	*n* = 3; 42.9%	"I would fall under that middle class, European descent, Christian-faith based person" (Christina**).
Treating others the way you want to be treated	*n* = 2; 28.6%	"I have values….treat people how you want to be treated" (Laura**).

Less Technology	*n* = 2; 28.6%	"I would say that would be my biggest difference is they're very much involved in technology where I am not. I did not grow up in that world. My culture is very not set in the use of technology" (Rebecca**).
Teamwork/Helping Others	*n* = 2; 28.6%	"Working together, collaboration, like that stuff that is really important to me" (Natalie).
Open-mindedness	*n* = 1; 14.3%	"There are times that I had been reminded, you know, like, 'Whoa, I better be careful, I better make sure that everybody actually celebrates that holiday, I better make sure that everybody actually feels that way'" (Julia).
Making mistakes	*n* = 1; 14.3%	"I really value mistakes and I really try to talk to [my students] about mistakes. So, I always say, 'what happens when you make a mistake?' and then they go, 'my brain grows bigger'" (Julia**).
Responsibility	*n* = 1; 14.3%	"I would say, responsibility is a big value of mine. Like, you do your homework before you do fun things. That type of thing was instilled" (Ava).
Hold eye contact	*n* = 1; 14.3%	"I have values….have eye contact" (Laura**).
Going to college	*n* = 1; 14.3%	"Going to college and having a work ethic" (Laura).
Honesty/trustworthiness	*n* = 1; 14.3%	"I suppose the biggest things that I value are hard work, honesty, respect, and, like, trustworthiness, I suppose" (Rebecca).

Note. A double asterisk (**) indicates the direct quote has been edited for clarity by removing filler words, linguistic utterances, and/or repeated words; however, the integrity of the quote was not changed.

Table 17

Research Question 7: Theme 5 -Types of Culturally Relevant Scenarios Teachers Would Add

Culturally Relevant Scenario Type	Number of Teacher Mentions (*n*); Percentage of Total Teachers (%)	Exemplar Quotes
Racial/ethnic specific scenarios	*n* = 4; 57.1%	**"**Maybe a few more examples from different cultures that they haven't experienced yet, like, I know there's a few in there. Like, Native Americans aren't really touched on and this area has a very huge population of Native Americans" (Rebecca**). "In terms of incorporating more Native American culture into school… I don't want to jump to like wild rice, but we just talked about this last week. But the scenario could be about a family that's going ricing or something. Like, just some scenario that has to do with a Native American tradition and kind of talk about that too" (Natalie).
Religious holidays specific scenarios	*n* = 2; 28.6%	"There could be a lesson just like, 'you might have someone who doesn't celebrate Christmas someday in your life, like how are you going to handle that?' You know? And it may take them till they're in high school before they meet someone who doesn't celebrate Christmas, but at least if we've talked about it every year since they were in kindergarten, they'll be like, 'oh yeah, this happens'" (Julia). "There are cultures that they aren't seeing their everyday life, but get them involved and like thinking about like, 'oh, other people might not celebrate like Christmas because everybody here celebrates Christmas.' And are...so getting them open to that possibility of like, Kwanza, and Hanukkah and that" (Rebecca**).

Variations in family-type scenarios	*n* = 1; 14.3%	"And even within the same ethnic backgrounds, like married versus divorced, parents, like that's a different culture, you know? Students that don't have their parents and are living with grandma and grandpa" (Laura).

Note. A double asterisk (**) indicates the direct quote has been edited for clarity by removing filler words, linguistic utterances, and/or repeated words; however, the integrity of the quote was not changed.

Table 18

Research Question 7: Theme 6 -Developmental Appropriateness Teachers Would Add

Developmentally Appropriate Action	Number of Teacher Mentions (*n*); Percentage of Total Teachers (%)	Exemplar Quotes
Engaging activities (e.g. movement based)	*n* = 2; 28.6%	"I would add more things to increase student engagement, just little activities to get the kids kind of moving a little more rather than sitting" (Susan). "I shortened down the scenarios at the end, and then I have kids come up with their own scenarios" (Natalie).
Simplify the steps	*n* = 2; 28.6%	"I would really simplify it. Like, you know, each skill we're trying to teach should have one or two steps, you know what I mean? Like, 'stop, think about it, do 'this' or something, not that every situation is…that black and white, but I think the first thing always needs to be 'stop' because first graders are generally very impulsive…it's got to be…way simpler" (Julia**).
More age appropriate videos/ Scenarios	*n* = 2; 28.6%	"Making some of the things a little more applicable for just third grade. But when you have a program that's K through [third grade]... that's not really realistic either…that's probably the biggest thing. Some of the videos are a little on the younger side" (Ava**)
Age-appropriate topics in the scenarios	*n* = 1; 14.3%	"Yeah, cuz [the scenario] was talking about pregnancy. Like, if somebody looks like they're having a baby or whatever. And I was like, 'Oh, this just doesn't seem appropriate for their age level.' Like, it's not something I would typically cover with them. Some of the examples weren't friendly for their age level" (Susan**).
More visuals	*n* = 1; 14.3%	" I would just make it more visuals less words. Because as soon as I start reading those, you know, too many sentences. It's just like right over their heads" (Ava**).

Shorten scenarios	*n* = 1; 14.3%	"The scenarios get to be a little bit like long, and I'd almost rather have like two or three scenarios and go into depth with them and really hash them out and do perspective taking and stuff like that with them versus having to go through six or seven short little things. Especially because fourth grade, they know the answers. Like, everybody knows the answers" (Natalie**).

Note. A double asterisk (**) indicates the direct quote has been edited for clarity by removing filler words, linguistic utterances, and/or repeated words; however, the integrity of the quote was not changed.

References

Abrams, Z. (2023). Teaching social-emotional learning is under attack. *Monitor on Psychology*, *54*(6), 28.https://www.apa.org/monitor/2023/09/social-emotional-learning-under-fire

Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine, 18*(3), 91-93. https://doi.org/10.1016/j.tjem.2018.08.001

Al Maskari, T. S., Melville, C. A., & Willis, D. S. (2018). Systematic review: Cultural adaptation and feasibility of screening for autism in non-English speaking countries. *International Journal of Mental Health Systems*, *12*(1), 22. https://doi.org/10.1186/s13033-018-0200-8

Alcántara, C., Giorgio Cosenzo, L., McCullough, E., Vogt, T., Falzon, A. L., & Perez Ibarra, I. (2021). Cultural adaptations of psychological interventions for prevalent sleep disorders and sleep disturbances: A systematic review of randomized controlled trials in the United States. *Sleep Medicine Reviews*, *56*, 101455. https://doi.org/10.1016/j.smrv.2021.101455

Algozzine, K., Christian, C., Marr, M. B., McClanahan, T., & White, R. (2008). Demography of problem behavior in elementary schools. *Exceptionality*, *16*(2), 93-104. https://doi.org/10.1080/09362830801981369

Allen, A., Hancock, S. D., Starker-Glass, T., & Lewis, C. W. (2017). Mapping culturally relevant pedagogy into teacher education programs: A critical framework. *Teacher College Record: The Voice of Scholarship in Education, 119*(1). https://doi-org.ezp3.lib.umn.edu/10.1177/01614681171190010

Allen, A., Scott, L. M., & Lewis, C. W. (2013). *Racial microaggressions and African American and Hispanic students in urban schools: A call for culturally affirming education*. *3*(2), 13. https://files.eric.ed.gov/fulltext/EJ1063228.pdf

Anderson, M. (2022). How social-emotional learning became a frontline in the battle against CRT. *National Public Radio*. https://www.npr.org/2022/09/26/1124082878/how-social-emotional-learning-became-a-front-line-in-the-battle-against-crt

Anzures, G., Quinn, P. C., Pascalis, O., Slater, A. M., & Lee, K. (2010). Categorization, categorical perception, and asymmetry in infants' representation of face race. *Developmental science*, *13*(4), 553-564. https://doi.org/10.1111/j.1467-7687.2009.00900.x

Arizona Department of Education. (2023, March 7). *Critical race theory and social emotional learning explained*. Arizona Department of Education. https://www.azed.gov/adeinfo/critical-race-theory-and-social-emotional-learning-explained

Bader, E. J. (2023). *The right is passing bills that ban the teaching of empathy and care in schools.* https://truthout.org/articles/the-right-is-passing-bills-that-ban-the-teaching-of-empathy-and-care-in-schools/

Bailey, R., Stickle, L., Brion-Meisels, G., & Jones, S. M. (2019). Re-imagining social-emotional learning: Findings from a strategy-based approach. *Phi Delta Kappan*, *100*(5), 53-58. https://doi-org.ezp1.lib.umn.edu/10.1177/0031721719827549

Baltes, P. B., Reese, H. W., & Lipsitt, L. P. (1980). Life-span developmental psychology. *Annual Review of Psychology*, *31*(1), 65–110. https://doi.org/10.1146/annurev.ps.31.020180.000433

Barrera, M., Castro, F. G., Strycker, L. A., & Toobert, D. J. (2013). Cultural adaptations of behavioral health interventions: A progress report. *Journal of Consulting and Clinical Psychology*, *81*(2), 196–205. https://doi.org/10.1037/a0027085

Baumann, A. A., Powell, B. J., Kohl, P. L., Tabak, R. G., Penalba, V., Proctor, E. K., Domenech-Rodriguez, M. M., & Cabassa, L. J. (2015). Cultural adaptation and implementation of evidence-based parent-training: A systematic review and critique of guiding evidence. *Children and Youth Services Review*, *53*, 113–120. https://doi.org/10.1016/j.childyouth.2015.03.025

Be Good People (2023). *Observation form*. https://drive.google.com/file/d/11t9LXp1fuddXUXhUG7igqDbUcwmctkN2/view

Be Good People (n.d). *Academic integration.* https://sites.google.com/scred.k12.mn.us/scredsel/educators/academic-integration

Bennet, D. A. (2001). How can I deal with missing data in my study? *Australian and New Zealand Journal of Public Health, 25*(5), 464-469. https://doi.org/10.1111/j.1467-842X.2001.tb00294.x

Bennouna, C., Brumbaum, H., McLay, M. M., Allaf, C., Wessells, M., & Stark, L. (2021). The role of culturally responsive social and emotional learning in supporting refugee inclusion and belonging: A thematic analysis of service provider perspectives. *Plos one*, *16*(8), e0256743. . https://doi.org/ 10.1371/journal.pone.0256743

Bernal, G., Jiménez-Chafey, M. I., & Domenech Rodríguez, M. M. (2009). Cultural adaptation of treatments: A resource for considering culture in evidence-based practice. *Professional Psychology: Research and Practice*, *40*, 361–368. https://doi.org/10.1037/a0016401

Blacksher, E., & Valles, S. A. (2021). White privilege, white poverty: Reckoning with class and race in America. *Hastings Center Report*, *51*, S51-S57. https://doi-org.ezp1.lib.umn.edu/10.1002/hast.1230

Bond, L. A., & Hauf, A. M. C. (2004). Taking stock and putting stock in primary prevention: Characteristics of effective programs. *Journal of Primary Prevention*, *24*, 199-221. https://link-springer-com.ezp3.lib.umn.edu/article/10.1023/B:JOPP.0000018051.90165.65

Bouffard, S. (2020). SEL and equity: 'If we don't address the wound, we cannot heal:' Q&A with Dena Simmons. *The Learning Professional, 41*(5), 24-27. https://learningforward.org/wp-content/uploads/2020/10/sel-and-equity-if-we-dont-address-the-wound-we-cannot-heal-QA-with-dena-simmons.pdf

Bowen, D. J., Kreuter, M., Spring, B., Cofta-Woerpel, L., Linnan, L., Weiner, D., ... & Fernandez, M. (2009). How we design feasibility studies. *American journal of preventive medicine*, *36*(5), 452-457. https://doi-org.10.1016/j.amepre.2009.02.002

Bowen, D. J., Kreuter, M., Spring, B., Cofta-Woerpel, L., Linnan, L., Weiner, D., Bakken, S., Kaplan, C. P., Squiers, L., Fabrizio, C., & Fernandez, M. (2009). How we design feasibility studies. *American Journal of Preventive Medicine, 36*(5), 452–457. https://doi.org/10.1016/j.amepre.2009.02.002

Brackett, M. A., Reyes, M. R., Rivers, S. E., Elbertson, N. A., & Salovey, P. (2012). Assessing teachers' beliefs about social and emotional learning. *Journal of Psychoeducational Assessment*, *30*(3), 219-236. https://doi.org/10.1177/0734282911424879

Braun, V., & Clarke, V. (2012). Thematic analysis. In *APA handbook of research methods in psychology, Vol 2: Research designs: Quantitative, qualitative, neuropsychological, and biological* (Vol. 2, pp. 57–71). American Psychological Association. https://doi.org/10.1037/13620-004

Briesch, A. M., Chafouleas, S. M., Neugebauer, S. R., & Riley-Tillman, T. C. (2013). Assessing influences on intervention implementation: Revision of the Usage Rating Profile-

Intervention. *Journal of school psychology*, *51*(1), 81-96. https://doi.org/10.1016/j.jsp.2012.08.006

Britton, W. B., Lepp, N. E., Niles, H. F., Rocha, T., Fisher, N. & Gold, J. (2015). A randomized controlled pilot trial of classroom-based mindfulness meditation compared to an active control condition in 6th grade children. *Journal of School Psychology*, 52(*3*), 263-278. https://doi.org/10.1016/j.jsp.2014.03.002

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Harvard University Press. https://books.google.com/books?id=OCmbzWka6xUC

Brown, C., Maggin, D. M., & Buren, M. (2018). Systematic review of cultural adaptations of school-based social, emotional, and behavioral interventions for students of color. *Education and Treatment of Children*, *41*(4), 431–456. https://doi.org/10.1353/etc.2018.0024

Brown-Jeffy, S., & Cooper, J. E. (2011). Toward a conceptual framework of culturally relevant pedagogy: An overview of the conceptual and theoretical literature. *Teacher Education Quarterly*, *38*(1), 65–84. https://www.jstor.org/stable/23479642

Bryan-Gooden, J., Hester, M., & Peoples, L. Q. (2023). *Culturally responsive ELA curriculum scorecard*. New York: Metropolitan Center for Research on Equity and the Transformation of Schools, New York University. https://steinhardt.nyu.edu/sites/default/files/2023-05/CRE%20ELA%20Curriculum%20Scorecard%202023.pdf

Buchanan, R., Gueldner, B. A., Tran, O. K., & Merrell, K. W. (2009). Social and Emotional Learning in Classrooms: A Survey of Teachers' Knowledge, Perceptions, and

Practices. *Journal of Applied School Psychology, 25*(2), 187–203. https://doi.org/10.1080/15377900802487078

Calgary Anti Racism Education. (2021). *Whiteness.* https://www.aclrc.com/whiteness

California Department of Education. (2023, March 9). *Culturally relevant pedagogy*. CA Dept of Education. https://www.cde.ca.gov/pd/ee/culturalrelevantpedagogy.asp

California Department of Education. (2024). *Evidence-based interventions under the ESSA.* California Department of Education. https://www.cde.ca.gov/re/es/evidence.asp

CASEL. (2023). *View All Programs—CASEL Program Guide*. CASEL. https://pg.casel.org/review-programs/

CASEL. (n.d.-a). *Fundamentals of SEL*. CASEL. Retrieved July 16, 2022, from https://casel.org/fundamentals-of-sel/

CASEL. (n.d.-b). *Our History*. CASEL. Retrieved May 12, 2023, from https://casel.org/about-us/our-history/

CASEL., (n.d.-c). *SEL Policy at the State Level.* Retrieved Aug 9, 2024 from https://casel.org/systemic-implementation/sel-policy-at-the-state-level/#MI

Castro Olivo, S. M. (2007). *The effects of a culturally -adapted social -emotional learning curriculum on social-emotional and academic outcomes of Latino immigrant high school students* [Ph.D., University of Oregon]. https://www.proquest.com/docview/304819586/abstract/1C299D5AD4E94037PQ/1

Castro Olivo, S. M., Ura, S., & dAbreu, A. (2022). The effects of a culturally adapted program on ELL students' core SEL competencies as measured by a modified version of the BERS-2. *Journal of Applied School Psychology*, *38*(4), 380–396. https://doi.org/10.1080/15377903.2021.1998278

Castro, F. G., Barrera, M., & Holleran Steiker, L. K. (2010). Issues and challenges in the design of culturally adapted evidence-based interventions. *Annual Review of Clinical Psychology*, *6*, 213–239. https://doi.org/10.1146/annurev-clinpsy-033109-132032

Castro-Olivo, S. M. (2014). Promoting social-emotional learning in adolescent Latino ells: A study of the culturally adapted Strong Teens program. *School Psychology Quarterly*, *29*(4), 567–577. https://doi.org/10.1037/spq0000055

Castro-Olivo, S. M., & Merrell, K. W. (2012). Validating cultural adaptations of a school-based social-emotional learning programme for use with Latino immigrant adolescents. *Advances in School Mental Health Promotion*, *5*(2), 78–92. https://doi.org/10.1080/1754730X.2012.689193

Chafouleas, S. M., Briesch, A. M., Neugebauer, S. R., & Riley-Tillman, T. C. (2011). *Usage Rating Profile – Intervention (Revised)*. Storrs, CT: University of Connecticut.

Chowdhary, N., Jotheeswaran, A. T., Nadkarni, A., Hollon, S. D., King, M., Jordans, M. J. D., Rahman, A., Verdeli, H., Araya, R., & Patel, V. (2014). The methods and outcomes of cultural adaptations of psychological treatments for depressive disorders: A systematic review. *Psychological Medicine*, *44*(6), 1131–1146. https://doi.org/10.1017/S0033291713001785

Cineas, F. (2023, February 13). *Conservatives' war on emotions in the classroom*. Vox. https://www.vox.com/the-highlight/23584837/social-emotional-learning-conservative-culture-war-in-schools

Cipriano, C., Strambler, M. J., Naples, L. H., Ha, C., Kirk, M., Wood, M., ... & Durlak, J. (2023). The state of evidence for social and emotional learning: A contemporary meta-analysis of universal school-based SEL interventions. *Child Development*, *94*(5), 1181-1204. https://doi.org/10.1111/cdev.13968

Clark, C., T., Chrisman, A., & Lewis, S. G. (2022). *(Un)standardizing emotions: An ethical critique of social and emotional learning standards*. *124*(7), 131–149. https://doi.org/10.1177/01614681221111432

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

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

Combi, M. (2016). Cultures and technology: An analysis of some of the changes in progress—Digital, global and local culture. In Borowiecki, K. J., Forbes, N., & Fresa, A. (Eds). *Cultural heritage in a changing world* (pp. 3-15). Springer Open. https://library.oapen.org/bitstream/handle/20.500.12657/28109/1001885.pdf?sequence=1#page=32

Committee for Children. (2011). *Second Step program*. https://www.secondstep.org/

Cooper, D. K., Wieling, E., Domenech Rodríguez, M. M., Garcia-Huidobro, D., Baumann, A., Mejia, A., Le, H.-N., Cardemil, E. V., & Acevedo-Polakovich, I. D. (2020). Latinx mental health scholars' experiences with cultural adaptation and implementation of systemic family interventions. *Family Process*, *59*(2), 492–508. https://doi.org/10.1111/famp.12433

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*(1), 98–104. https://doi-org.ezp3.lib.umn.edu/10.1037/0021-9010.78.1.98

Cramer, K., & Castro-Olivo, S. (2015). Effects of a culturally adapted social-emotional learning intervention program on students' mental health. *Contemporary School Psychology (Springer Science & Business Media B.V.)*, *20*(2), 118–129. https://doi.org/10.1007/s40688-015-0057-7

Credé, M. (2018). What shall we do about grit? A critical review of what we know and what we don't know. *Educational Researcher*, *47*(9), 606–611. https://doi.org/10.3102/0013189X18801322

Cuocci, S., & Arndt, R. (2020). SEL for culturally and linguistically diverse students. *Journal of English Learner Education*, *10*(1). https://stars.library.ucf.edu/jele/vol10/iss1/4

Dancey, C. P., & Reidy, J. (2007). *Statistics without maths for psychology*. Pearson education

Darling-Hammond, L., Schachner, A., & Edgerton, A. K. (2020). Restarting and Reinventing School: Learning in the Time of COVID and Beyond. *Learning Policy Institute*. https://eric.ed.gov/?id=ED610890

Datu, J. A. D., & Restubog, S. L. D. (2020). The emotional pay-off of staying gritty: Linking grit with social-emotional learning and emotional well-being. *British Journal of Guidance & Counselling*, *48*(5), 697–708. https://doi.org/10.1080/03069885.2020.1758922

Davis, C. U. (2018). Laying new ground. *Journal of Dance Education*, *18*(3), 120–125. https://doi.org/10.1080/15290824.2018.1481965

Denham, S. A. (2018). Keeping SEL developmental: The importance of a developmental lens for fostering and assessing SEL competencies. *Frameworks Briefs*, *1*(2), 127-152.https://www.militarychild.org/upload/images/MGS%202022/WellbeingToolkit/PDFs/2_EI_Keeping-SEL-Developmental-CASEL.pdf

Devakumar, D., Selvarajah, S., Shannon, G., Muraya, K., Lasoye, S., Corona, S., Paradies, Y., Abubakar, I., & Achiume, E. T. (2020). Racism, the public health crisis we can no longer ignore. *The Lancet*, *395*(10242), e112–e113. https://doi.org/10.1016/S0140-6736(20)31371-4

Drake University. (2023). The Six Pillars of Character. *Character Counts*. https://charactercounts.org/character-counts-overview/six-pillars/

Drug Enforcement Administration. (2023). *Why do teens use drugs?* https://www.getsmartaboutdrugs.gov/family/why-do-teens-use-drugs

Duncan, R., Washburn, I. J., Lewis, K. M., Bavarian, N., DuBois, D. L., Acock, A. C., ... & Flay, B. R. (2016). Can universal SEL programs benefit universally? Effects of the positive action program on multiple trajectories of social-emotional and misconduct behaviors. *Prevention science*, *18*(2), 214-224. https://doi.org/10.1007/s11121-016-0745-1

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*(1), 405–432. https://doi.org/10.1111/j.1467-8624.2010.01564.x

Dusenbury, L., & Weissberg, R. P. (2017). *State efforts to promote social and emotional learning in students*. https://files.eric.ed.gov/fulltext/ED581609.pdf

Eklund, K., Kilpatrick, K. D., Kilgus, S. P., & Haider, A. (2018). A systematic review of state-level social–emotional learning standards: Implications for practice and research. *School Psychology Review*, *47*(3), 316–326. https://doi.org/10.17105/SPR-2017.0116.V47-3

Elbertson, N. A., Brackett, M. A., & Weissberg, R. P. (2010). School-based social and emotional learning (sel) programming: Current perspectives. In A. Hargreaves, A. Lieberman, M. Fullan, & D. Hopkins (Eds.), *Second International Handbook of Educational Change* (pp. 1017–1032). Springer Netherlands. https://doi.org/10.1007/978-90-481-2660-6_57

Elek, E., Wagstaff, D. A., & Hecht, M. L. (2010). Effects of the 5th and 7th grade enhanced versions of the keepin' it REAL substance use prevention curriculum. *Journal of Drug Education*, *40*(1), 61–79. https://doi.org/10.2190/DE.40.1.e

Elliot, S. N., Davies, M. D., Frey, J. R., Gresham, F., Cooper, G. (2018). Development and initial validation of a social emotional learning assessment for universal screening. *Journal of Applied Developmental Psychology, 55*, 39-51. https://doi.org/10.1016/j.appdev.2017.06.002

Evans, O. G. (2023). Bronfenbrenner's ecological systems theory. *Simply Psychology*. https://www.simplypsychology.org/bronfenbrenner.html

Fahle, E. M., Reardon, S. F., Kalogrides, D., Weathers, E. S., & Jang, H. (2020). Racial segregation and school poverty in the United States, 1999–2016. *Race and Social Problems*, *12*(1), 42–56. https://doi.org/10.1007/s12552-019-09277-w

Fastbridge. (2023-a). mySAEBRS overview. *Renaissance.* https://fastbridge.illuminateed.com/hc/en-us/articles/1260802484089-mySAEBRS-Overview.

Fastbridge. (2023-b). SAEBRS and mySAEBRS norms and benchmarks. *Renaissance.* https://fastbridge.illuminateed.com/hc/en-us/articles/1260802344370-SAEBRS-and-mySAEBRS-Norms-and-Benchmarks

Fastbridge. (2023-c). SAEBRS overview. *Renaissance.* https://fastbridge.illuminateed.com/hc/en-us/articles/1260802463290-SAEBRS-Overview

Fenton, L., Lauckner, H., & Gilbert, R. (2015). The QATSDD critical appraisal tool: Comments and critiques. *Journal of Evaluation in Clinical Practice*, *21*(6), 1125–1128. https://doi.org/10.1111/jep.12487

Ferrer-Wreder, L., Eninger, L., Olsson, T. M., Sedem, M., Allodi, M. W., & Ginner Hau, H. (2021). The cultural adaptation of interventions to promote positive development: The preschool edition of PATHS® in Sweden. In R. Dimitrova & N. Wiium (Eds.), *Handbook of Positive Youth Development: Advancing Research, Policy, and Practice*

in Global Contexts (pp. 399–413). Springer International Publishing. https://doi.org/10.1007/978-3-030-70262-5_27

Fineran, S. (2002). Sexual harassment between same-sex peers: Intersection of mental health, homophobia, and sexual violence in schools. *Social Work*, *47*(1), 65–74. https://doi.org/10.1093/sw/47.1.65

Fisher, H. H., Hawkins, G. T., Hertz, M., Sliwa, S., & Beresovsky, V. (2022). Student and School Characteristics Associated With COVID-19-Related Learning Decline Among Middle and High School Students in K-12 Schools. *Journal of School Health*, *92*(11), 1027-1039. https://doi-org.ezp1.lib.umn.edu/10.1111/josh.13243

Fiske, A. P. (2002). Using individualism and collectivism to compare cultures--A critique of the validity and measurement of the constructs: Comment on Oyserman et al. (2002). *Psychological Bulletin*, *128*, 78–88. https://doi.org/10.1037/0033-2909.128.1.78

Foste, Z., & Tevis, T. L. (Eds.). (2023). *Critical whiteness praxis in higher education: Considerations for the pursuit of racial justice on campus*. Taylor & Francis. https://books.google.com/books?hl=en&lr=&id=rhHJEAAAQBAJ&oi=fnd&pg=PA2016&dq=tevis+and+foste+2023&ots=pTe5fQi_ww&sig=cQBfRlI8g2oXo58k2L_JyNa38h4#v=onepage&q=tevis%20and%20foste%202023&f=false

Foster, S. L., & Mash, E. J. (1999). Assessing social validity in clinical treatment research: Issues and procedures. *Journal of Consulting and Clinical Psychology*, *67*(3), 308. https://doi.org/10.1037/0022-006X.67.3.308

Freire, P. (1970). *Pedagogy of the oppressed: 50th anniversary edition*. Bloomsbury Publishing USA. https://books.google.com/books?id=OrVLDwAAQBAJ

Gay, G. (1994). *At the essence of learning: Multicultural education*. Kappa Delta Pi, 1601 West State Street, Post Office Box A, West Lafayette, IN 47906-0576 ($15). https://eric.ed.gov/?id=ED386410

Gay, G. (2002). Preparing for culturally responsive teaching. *Journal of Teacher Education, 53*(2), 106-116. https://doi.org/10.1177/0022487102053002003

Gay, G. (2018). *Culturally responsive teaching: Theory, researcher, and practice* (3rd ed.). Teachers College Press. https://books.google.com/books?hl=en&lr=&id=uD9qDwAAQBAJ&oi=fnd&pg=PP1&dq=gay+2000+culturally+responsive+teaching&ots=v0tC760ieb&sig=BRYCFyFxJriFQcsN6NOwmFIN3To#v=onepage&q=gay%202000%20culturally%20responsive%20teaching&f=false

Goerdt, A. (2023). Technical adequacy of the social, academic, and emotional behavior risk screener - teacher rating scale: A systematic review, quantitative synthesis, & measurement invariance study. Retrieved from the University Digital Conservancy, https://hdl.handle.net/11299/256996

Gosin, M., Marsiglia, F. F., & Hecht, M. L. (2003). Keepin' it R.E.A.L.: A drug resistance curriculum tailored to the strengths and needs of pre-adolescents of the southwest. *Journal of Drug Education*, *33*(2), 119–142. https://doi.org/10.2190/DXB9-1V2P-C27J-V69V

Graves, S. L. Jr., & Aston, C. (2018). A mixed-methods study of a social emotional curriculum for Black male success: A school-based pilot study of the Brothers of Ujima. *Psychology in the Schools*, *55*(1), 76–84. https://doi.org/10.1002/pits.22088

Graves, S. L., Jr., Herndon-Sobalvarro, A., Nichols, K., Aston, C., Ryan, A., Blefari, A., Schutte, K., Schachner, A., Vicoria, L., & Prier, D. (2017). Examining the effectiveness of a culturally adapted social-emotional intervention for African American males in an

urban setting. *School Psychology Quarterly*, *32*(1), 62–74. https://doi.org/10.1037/spq0000145

Greenberg, M. T., Domitrovich, C. E., Weissberg, R. P., & Durlak, J. A. (2017). Social and emotional learning as a public health approach to education. *The Future of Children*, *27*(1), 13–32. https://doi.org/10.1353/foc.2017.0001

Gregory, A., & Fergus, E. (2017). Social and emotional learning and equity in school discipline. *The Future of Children*, *27*(1), 117–136. https://www.jstor.org/stable/44219024

Gresham, F. M. (1996). *SSIS Social Emotional Learning Edition.* https://www.pearsonassessments.com/store/usassessments/en/Store/Professional-Assessments/Behavior/SSIS-Social-Emotional-Learning-Edition/p/100001940.html

Gresham, F., & Elliott, S. N. (2008). *Social skills improvement system (SSIS) rating scales.* Bloomington, MN: Pearson Assessments.

Griner, D., & Smith, T. B. (2006). Culturally adapted mental health intervention: A meta-analytic review. *Psychotherapy (Chicago, Ill.)*, *43*(4), 531–548. https://doi.org/10.1037/0033-3204.43.4.531

Hall, H., & Nielsen, E. (202, May 26). *How do children spend their time? Time use and skill development in the PSID.* Board of Governors of the Federal Reserve System. https://www.federalreserve.gov/econres/notes/feds-notes/how-do-children-spend-their-time-time-use-and-skill-development-in-the-psid-20200526.html#:~:text=Not%20surprisingly%2C%20sleep%20is%20the,personal%20hygiene%2C%20at%2015%20percent.

Harrison, R., Jones, B., Gardner, P., & Lawton, R. (2021). Quality assessment with diverse studies (QuADS): An appraisal tool for methodological and reporting quality in

systematic reviews of mixed- or multi-method studies. *BMC Health Services Research*, *21*(1), 144. https://doi.org/10.1186/s12913-021-06122-y

Hatchimonji, D. R. (2016). *Grit in Latino middle school students: Construct validity and psychometric properties of the Short Grit Scale* [M.S., Rutgers The State University of New Jersey, School of Graduate Studies]. https://www.proquest.com/docview/1781032723/abstract/1949017528FD4807PQ/1

Hayes, C., & Juarez, B. (2011). There is no culturally responsive teaching spoken here: A critical race perspective. *Democracy and Education, 20*(1). ttps://democracyeducationjournal.org/home/vol20/iss1/1

Hecht, M. L., Shin, Y., Pettigrew, J., Miller-Day, M., & Krieger, J. L. (2018). Designed cultural adaptation and delivery quality in rural substance use prevention: An effectiveness trial for the keepin' it REAL curriculum. *Prevention Science*, *19*(8), 1008–1018. https://doi.org/10.1007/s11121-018-0937-y

Ho, J., & Funk, S. (2018). Preschool: Promoting young children's social and emotional health. *YC Young Children*, *73*(1), 73–79. https://www.jstor.org/stable/90019485

Hoffman, D. M. (2009). Reflecting on social emotional learning: A critical perspective on trends in the united states. *Review of Educational Research*, *79*(2), 533–556. https://doi.org/10.3102/0034654308325184

Hopson, L. M. (2006). *Effectiveness of culturally grounded adaptations of an evidence-based substance abuse prevention program with alternative school students* [Ph.D., The University of Texas at Austin]. https://www.proquest.com/docview/304983580/abstract/98C5DCC5180747DFPQ/1

Imada, T. (2010). Cultural narratives of individualism and collectivism: A content analysis of textbook stories in the United States and Japan. *Journal of Cross-Cultural Psychology*, *43*(4), 576–591. https://doi.org/10.1177/0022022110383312

Jagers, R. J., Rivas-Drake, D., & Borowski, T. (2018). Equity & social and emotional learning: A cultural analysis. *CASEL Assessment Work Group Brief series*. https://drc.casel.org/uploads/sites/3/2019/02/Equity-Social-and-Emotional-Learning-A-Cultural-Analysis.pdf

Jagers, R. J., Rivas-Drake, D., & Williams, B. (2019). Transformative social and emotional learning (SELl): Toward sel in service of educational equity and excellence. *Educational Psychologist*, *54*(3), 162–184. https://doi.org/10.1080/00461520.2019.1623032

Jagers, R. J., Skoog-Hoffman, A., Barthelus, B., & Schlund, J. (2021). Transformative social emotional learning: In pursuit of educational equity and excellence. *American Educator*, *45*(2), 12. https://eric.ed.gov/?id=EJ1304336

Johnstone, J., Ribbers, A., Jenkins, D., Atchley, R., Gustafsson, H. ,Nigg, J., Wahbeh, H., & Oken, B. (2020). Classroom-based mindfulness training reduces anxiety in adolescents: Acceptability and effectiveness of a cluster-randomized pilot study.

Jones, S. M., & Doolittle, E. J. (2017). Social and emotional learning: Introducing the issue. *The Future of Children*, *27*(1), 3–11. https://www.jstor.org/stable/44219018

Kaler-Jones, C. (2020, May 7). When SEL is used as another form of policing. *Medium*. https://medium.com/@justschools/when-sel-is-used-as-another-form-of-policing-fa53cf85dce4

Kendziora, K., & Yoder, N. (2016). When Districts Support and Integrate Social and Emotional Learning (SEL): Findings from an Ongoing Evaluation of Districtwide

Implementation of SEL. *Education Policy Center at American Institutes for Research*. https://doi-org.ezp1.lib.umn.edu/10.1177/0031721719827549

Kilgus, S. P., Eklund, K., von der Embse, N. P., Weist, M., Barber, A. J., Kaul, M., & Dodge, S. (2020). Structural validitgy and reliability of social, academic, and emotional behavior risk screener-student rating scales: A replication study. *Assessment for Effective Intervention, 46*(4). https://doi-org.ezp3.lib.umn.edu/10.1177/1534508420909527

King, G. (2023, February 25). *Iowa bill could ban social-emotional learning*. https://www.thegazette.com/k/iowa-bill-seeks-to-ban-social-emotional-learning-but-educators-say-its-critical-to-kids-developme/

Klees, S. J. (2020). Beyond neoliberalism: Reflections on capitalism and education. *Policy Futures in Education*, *18*(1), 9–29. https://doi.org/10.1177/1478210317715814

Kreuter, M. W., & Wray, R. J. (2003). Tailored and targeted health communication: Strategies for enhancing information relevance. *American Journal of Health Behavior*, *27*(1), S227–S232. https://doi.org/10.5993/AJHB.27.1.s3.6

Kurtz, K. D. (2020). *Adapting social emotional learning to impact internalizing problems in an urban school* [Ph.D., University of Massachusetts Boston]. https://www.proquest.com/docview/2417988419/abstract/BAADC25B47524D61PQ/1

Kwok, M., Rios, A., & Kwok, A. (2023). Dispelling deficit framing: Investigating preservice teacher beliefs about the intersections between literacy instruction and culturally relevant pedagogy. *Teachers and Teaching, 29*(1), 52-69. https://doi-org.ezp3.lib.umn.edu/10.1080/13540602.2022.2144821

Labaree, D. F. (2009). Educational formalism and the language of goals in american education, educational reform, and educational history. *Educational Research:*

Proofs, Arguments, and Other Reasonings, 41. https://doi.org/10.1007/978-90-481-3249-2_4

LaBelle, B. (2019). Positive outcomes of a social-emotional learning program to promote student resiliency and address mental health. *Contemporary School Psychology*, *27*(1), 1-7. https://doi.org/10.1007/s40688-019-00263-y

Ladson-Billings, G. (1994). *The dreamkeepers. Successful teachers of African American children*. Jossey-Bass Inc. https://eric.ed.gov/?id=ED373128

Ladson-Billings, G. (1995). Toward a theory of culturally relevant pedagogy. *American Educational Research Journal*, *32*(3), 465–491. https://doi.org/10.2307/1163320

Ladson-Billings, G. (2003). Lies my teacher still tells: Developing a critical race perspective toward the social studies. In Ladson-Billings, G. (Eds). *Critical race theory perspectives on social studies* (pp. 1-11). Information Age Publishing. https://books.google.com/books?hl=en&lr=&id=lPgnDwAAQBAJ&oi=fnd&pg=PA1&dq=things+my+history+teacher+never+taught+me&ots=n3RKLrVi-A&sig=FYB5tUb_wpuwjgL7roy4YQoOOTM#v=onepage&q&f=false

Ladson-Billings, G., & Tate, W. F. (1995). Toward a critical race theory of education. *Teachers College Record*, *97*(1), 47–68. https://www.proquest.com/docview/211325429

Leaper, C., & Spears Brown, C. (2014). Chapter six: Sexism in schools. In L. S. Liben & R. S. Bigler (Eds.), *Advances in Child Development and Behavior* (Vol. 47, pp. 189–223). JAI. https://doi.org/10.1016/bs.acdb.2014.04.001

Lin, M., Olsen, S., Simmons, D. N., Miller, M., & Tominey, S. L. (2023, January). "Not try to save them or ask them to breathe through their oppression": Educator perceptions and the need for a human-centered, liberatory approach to social and emotional

learning. In *Frontiers in Education* (Vol. 7, p. 1044730). Frontiers Media SA. https://doi.org/10.3389/feduc.2022.1044730

López, Q. (2022, October 10). *Latinx, Latine, or Latino? 8 Queer and Trans People tell us what they prefer and why*. Them. https://www.them.us/story/latinx-latine-difference-definition

Losen, D. J., Hodson, C. L., Keith II, M. A., Morrison, K., & Belway, S. (2015). *Are we closing the school discipline gap?* https://escholarship.org/uc/item/2t36g571

Love, B. L. (2019, February 13). "Grit is in our DNA": Why teaching grit is inherently anti-Black. *Education Week*. https://www.edweek.org/leadership/opinion-grit-is-in-our-dna-why-teaching-grit-is-inherently-anti-black/2019/02

Luevano, C., & Collins, T. A. (2020). Culturally appropriate math problem-solving instruction with English language learners. *School Psychology Review*, *49*(2), 144–160. https://doi.org/10.1080/2372966X.2020.1717243

Mahfouz, J., & Anthony-Stevens, V. (2020). Why trouble SEL? The need for cultural relevance in SEL. *Occasional Paper Series*, *2020*(43). https://doi.org/10.58295/2375-3668.1354

Mandalaywala, T. M., Tai, C., & Rhodes, M. (2020). Children's use of race and gender as cues to social status. *PloS one*, *15*(6), e0234398. https://doi.org/10.1371/journal.pone.0234398

Marshall, S., Taki, S., Laird, Y., Love, P., Wen, L. M., & Rissel, C. (2021). Cultural adaptations of obesity-related behavioral prevention interventions in early childhood: A systematic review. *Obesity Reviews*, *23*(4), e13402. https://doi.org/10.1111/obr.13402

Marsiglia, F. F., Kulis, S. S., Cutrín, O., Medina-Mora, M. E., Real, T., Nuño-Gutiérrez, B. L., Corona, M. D., Mendoza-Meléndez, M. Á., Gresenz, K., & Alcala-Calvillo, D. (2022).

The feasibility, acceptability, and utility of Mantente REAL: The culturally adapted version of keepin' it REAl for Mexico. *Prevention Science*, *23*(8), 1483–1494. https://doi.org/10.1007/s11121-022-01409-w

Martin, C. L., & Ruble, D. (2004). Children's search for gender cues: Cognitive perspectives on gender development. *Current Directions in Psychological Science*, *13*(2), 67–70. https://doi.org/10.1111/j.0963-7214.2004.00276.x

Martinez, C. S. (2011). *Spark: Program evaluation of school-based group therapy for at-risk students* [Psy.D., The Wright Institute]. https://www.proquest.com/docview/875292521/abstract/6F39D47C99D34633PQ/1

Martínez, L. (2016). Teachers' voices on social emotional learning : identifying the conditions that make implementation possible. *The International Journal of Emotional Education, 8*(2), 6-24. https://www.um.edu.mt/library/oar/handle/123456789/14429

Matsuba, M. K., Schonert-Reichl, K. A., McElroy, T., & Katahoire, A. (2021). Effectiveness of a SEL/mindfulness program on Northern Ugandan children. *International Journal of School & Educational Psychology*, *9*(sup1), S113–S131. https://doi.org/10.1080/21683603.2020.1760977

McCallops, K., Barnes, T. N., Berte, I., Fenniman, J., Jones, I., Navon, R., & Nelson, M. (2019). Incorporating culturally responsive pedagogy within social-emotional learning interventions in urban schools: An international systematic review. *International Journal of Educational Research*, *94*, 11–28. https://doi.org/10.1016/j.ijer.2019.02.007

McDougal, J. L., Bardos, A. N., & Meier, S. T. (2011). Behavior Intervention Monitoring Assessment System Technical Manual. Toronto: Multi-Health Systems.

https://www.rand.org/education-and-labor/projects/assessments/tool/1998/behavior-intervention-monitoring-assesment-system-teacher.html

Mcleod, S. (2023, April 14). *Social identity theory: Definition, history, examples, & facts*. https://www.simplypsychology.org/social-identity-theory.html

Mendez, D. D., Scott, J., Adodoadji, L., Toval, C., McNeil, M., & Sindhu, M. (2021). Racism as a public health crisis: Assessment and review of municpal declarations and resolutions across the United States. *Frontiers Public Health, 9.* https://doi.org/10.3389/fpubh.2021.686807

Merrell, K. W., Carrizales, D. C., Feuerborn, L., Gueldner, B. A. and Tran, O. K. 2007a. Strong Kids—grades 3–5: A social-emotional learning curriculum, Baltimore: Paul H. Brookes Publishing.

Merrell, K. W., Carrizales, D. C., Feuerborn, L., Gueldner, B. A. and Tran, O. K. 2007b. Strong Kids—grades 6–8: A social-emotional learning curriculum, Baltimore: Paul H. Brookes Publishing.

Merrell, K. W., Carrizales, D. C., Feuerborn, L., Gueldner, B. A. and Tran, O. K. 2007c. Strong Teens—grades 9–12: A social-emotional learning curriculum, Baltimore: Paul H. Brookes Publishing.

Midgett, A., Doumas, D., Sears, D., Lundquist, A., & Hausheer, R. (2015). A bystander bullying psychoeducation program with middle school students: A preliminary report. *The Professional Counselor*. http://dx.doi.org/10.15241/am.5.4.486

Mier, N., Ory, M. G., & Medina, A. A. (2010). Anatomy of culturally sensitive interventions promoting nutrition and exercise in Hispanics: A critical examination of existing literature. *Health Promotion Practice*, *11*(4), 541–554. https://doi.org/10.1177/1524839908328991

Miller, T., Volmert, A., Rochman, A., & Aassar, M. (2021). Talking about poverty: Narratives, counter-narratives, and telling effective stories. *FrameWorks Institute*. https://www.frameworksinstitute.org/publication/talking-about-poverty-narratives-counter-narratives-and-telling-effective-stories/

Milner, H. R., & Laughter, J. C. (2015). But good intentions are not enough: Preparing teachers to center race and poverty. *The Urban Review, 47*, 341-363. https://link-springer-com.ezp3.lib.umn.edu/article/10.1007/s11256-014-0295-4

Minnesota Department of Education. (2024). *SEL implementation guide.* Minnesota Department of Education. https://education.mn.gov/MDE/dse/safe/social/imp/

Minnesota Report Card. (2023). *Search to find your school.* https://rc.education.mn.gov/#mySchool/p--3

Monroe, C. R. (2005). Understanding the discipline gap through a cultural lens: Implications for the education of African American students. *Intercultural Education*, *16*(4), 317–330. https://doi.org/10.1080/14675980500303795

Moran, M., Midgett, A., Doumas, D. M., Porchia, S., & Moody, S. (2019). A mixed method evaluation of a culturally adapted, brief, bullying bystander intervention for middle school students. *Journal of Child and Adolescent Counseling*, *5*(3), 221–238. https://doi.org/10.1080/23727810.2019.1669372

Nagayama Hall, G. C., Ibaraki, A. Y., Huang, E. R., Marti, C. N., & Stice, E. (2016). A meta-analysis of cultural adaptations of psychological interventions. *Behavior Therapy*, *47*(6), 993–1014. https://doi.org/10.1016/j.beth.2016.09.005

National Center for Education Statistics. (2022). *More than 80 percent of U.S. public schools report pandemic has negatively impacted student behavior and socio-emotional development.* https://nces.ed.gov/whatsnew/press_releases/07_06_2022.asp

National Center for Education Statistics. (2022-b). *Racial/Ethnic Enrollment in Public Schools*. IES. https://nces.ed.gov/programs/coe/indicator/cge/racial-ethnic-enrollment

National Center for Education Statistics. (2023). *Racial/ethnic enrollment in public schools.* https://nces.ed.gov/programs/coe/indicator/cge/racial-ethnic-enrollment

National Center for Education Statistics. (2023-a). *Characteristics of public school teachers.* https://nces.ed.gov/programs/coe/indicator/clr/public-school-teachers

National Center for Education Statistics. (2023-b). *Race/ethnicity of college faculty.* https://nces.ed.gov/fastfacts/display.asp?id=61

National Center or Education Statistics. (2023-c). *Racial/ethnic enrollment of public schools.* https://nces.ed.gov/programs/coe/indicator/cge/racial-ethnic-enrollment

National Museum of African American History and Culture. (n.d). *Whiteness.* Smithsonian. https://nmaahc.si.edu/learn/talking-about-race/topics/whiteness

Netemeyer, R. G., Bearden, W. O., & Sharma, S. (2003). Scaling procedures: Issues and applications. sage publications.

Nguyen, B. M. D., Noguera, P., Adkins, N., & Teranishi, R. T. (2019). Ethnic discipline gap: Unseen dimensions of racial disproportionality in school discipline. *American Educational Research Journal*, *56*(5), 1973–2003. https://doi.org/10.3102/0002831219833919

Nierkens, V., Hartman, M. A., Nicolaou, M., Vissenberg, C., Beune, E. J. A. J., Hosper, K., Valkengoed, I. G. van, & Stronks, K. (2013). Effectiveness of cultural adaptations of interventions aimed at smoking cessation, diet, and/or physical activity in ethnic minorities. A systematic review. *PLOS ONE*, *8*(10), e73373. https://doi.org/10.1371/journal.pone.0073373

Nieto, S. (1999). *The light in their eyes: Creating multicultural learning communities.* Teachers College Press, 1234 Amsterdam Avenue, New York, NY 10027 (paperback: ISBN-0-8077-3782-8, $22. https://eric.ed.gov/?id=ED440164

Ogbu, J. U. (1987). Variability in minority school performance: A problem in search of an explanation. *Anthropology & Education Quarterly*, *18*(4), 312–334. https://www.jstor.org/stable/3216660

Otter.ai (2024). *Otter.ai*. https://otter.ai/

Owens, J. S., Holdaway, A. S., Smith, J., Evans, S. W., Himawan, L. K., Coles, E. K., ... & Dawson, A. E. (2018). Rates of common classroom behavior management strategies and their associations with challenging student behavior in elementary school. *Journal of Emotional and Behavioral Disorders*, *26*(3), 156-169.https://doi-org.ezp3.lib.umn.edu/10.1177/1063426617712

Paris, D. (2012). Culturally sustaining pedagogy: A needed change in stance, terminology, and practice. Educational Researcher, 41(3), 93–97. https://doi.org/10.3102/0013189X12441244

Payton, J., Weissberg, R. P., Durlak, J. A., Dymnicki, A. B., Taylor, R. D., Schellinger, K. B., & Pachan, M. (2008). The positive impact of social and emotional learning for kindergarten to eighth-grade students: Findings from three scientific reviews. In *Collaborative for Academic, Social, and Emotional Learning (NJ1)*. Collaborative for Academic, Social, and Emotional Learning. https://eric.ed.gov/?id=ED505370

Perepletchikova, F., & Kazdin, A. E. (2005). Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical Psychology: Science and Practice, 12*(4), 365–383. https://doi-org.ezp3.lib.umn.edu/10.1093/clipsy.bpi045

Pietrantoni, Z., Glance, D., & Smith, K. M. (2015). Experiences with classism: A look at social class in a rural high school. *Journal of School Counseling*, *13*(12). https://eric.ed.gov/?id=EJ1069403

Podoshen, J. S., & Ekpo, A. E. (2021). Diversity, tokenism, and comic books: Crafting better strategies. *Business Horizons*, *64*(1), 131-140. https://doi.org/10.1016/j.bushor.2020.10.006

Qayyum, A., & Hussain, S. (2019). *Effect of scaffolding on social and emotional (SEL) of elementary school students*. *27*(1), 83–92. https://ojs.aiou.edu.pk/index.php/jssh/article/download/601/532

Quinn, P. C., Anzures, G., Izard, C. E., Lee, K., Pascalis, O., Slater, A. M., & Tanaka, J. W. (2011). Looking across domains to understand infant representation of emotion. *Emotion Review*, *3*(2), 197–206. https://doi.org/10.1177/1754073910387941

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Ramirez, T., Brush, K., Raisch, N., Bailey, R., & Jones, S. M. (2021). Equity in social emotional learning programs: A content analysis of equitable practices in prek-5 SEL programs. *Frontiers in Education*, *6*. https://doi.org/10.3389/feduc.2021.679467

Reed, M. (2022). *Should we banish the word "stakeholder"?* Fast Track Impact. https://www.fasttrackimpact.com/post/why-we-shouldn-t-banish-the-word-stakeholder

Regional Educational Laboratory (2019). *ESSA tiers of evidence: What you need to know.* https://ies.ed.gov/ncee/edlabs/regions/midwest/pdf/blogs/RELMW-ESSA-Tiers-Video-Handout-508.pdf

Resnicow, K., Baranowski, T., Ahluwalia, J., & Braithwaite, R. (1999). Cultural sensitivity in public health: Defined and demystified. *Ethnicity & Disease*, *9*, 10–21. https://pubmed-ncbi-nlm-nih-gov.ezp1.lib.umn.edu/10355471/

Ryan, A. M. (2002). Core Elements of Preparing Teachers for Culturally Relevant Practice.https://eric.ed.gov/?id=ED467795

Schaeffer, K. (2021). *America's public school teachers are far less racially and ethnically diverse than their students.* Pew Research Center. https://www.pewresearch.org/short-reads/2021/12/10/americas-public-school-teachers-are-far-less-racially-and-ethnically-diverse-than-their-students/

Schwartz, H. L., Bongrad, M., Bogan, E. D., Boyle, A. E., Meyers, D. C., & Jagers, R. J. (2022). *Social and emotional learning in schools nationally and in the collaborating districts initiative: Selected findings from the American teacher panel and American school leader panel surveys*. RAND Corporation. https://doi.org/10.7249/RRA1822-1

Shapiro, S. (2014). "Words that you said got bigger": English language learners' lived experiences of deficit discourse. Research in the Teaching of English, 48(4), 386-406.https://doi.org/10.58680/rte201425159

Sharma, S. A., & Christ, T. (2017). Five steps toward successful culturally relevant text selection and integration. *The Reading Teacher*, *71*(3), 295-307.https://doi.org/10.1002/trtr.1623

Shutts, K. (2015). Young children's preferences: Gender, race, and social status. *Child Development Perspectives*, *9*(4), 262–266. https://doi.org/10.1111/cdep.12154

Sirriyeh, R., Lawton, R., Gardner, P., & Armitage, G. (2011). Reviewing studies with diverse designs: The development and evaluation of a new tool. *Journal of Evaluation in Clinical Practice*, *18*(4), 746–752. https://doi.org/10.1111/j.1365-2753.2011.01662.x

Sivaraman, M., & Fahmie, T. A. (2020). Evaluating the efficacy and social validity of a culturally adapted training program for parents and service providers in India. *Behavior Analysis in Practice*, *13*(4), 849–861. https://doi.org/10.1007/s40617-020-00489-w

Six Seconds. (n.d.). *Six seconds: The emotional intelligence network.* https://www.6seconds.org/

Slaten, C. D., Irby, D. J., Tate, K., & Rivera, R. (2015). Towards a critically conscious approach to social and emotional learning in urban alternative education: School staff members' perspectives. *Journal for Social Action in Counseling & Psychology*, *7*(1), Article 1. https://doi.org/10.33043/JSACP.7.1.41-62

Smolicz, J. (1981). Core values and cultural identity. *Ethnic and Racial studies, 4*(1), 75-90. https://doi.org/10.58680/rte201425159

Spencer-Oatey, H., & Franklin, P. (2012). What is culture. *A compilation of quotations. GlobalPAD Core Concepts*, *1*(22), 1-21. https://www.academia.edu/download/40397043/Spencer-Oatey_2012_What_is_Culture.pdf

St. Croix River Education District. (2022, August 22-a). *Creating a school-wide schedule for SEL instruction (elementary schools).* [Video]. YouTube. https://www.youtube.com/watch?v=i_SOt67oUEw&t=1s

St. Croix River Education District. (2022, August 25-b). *A teacher's guide to the Be Good People SEL curriculum: Core edition, grades K-3.* [Video]. YouTube. https://www.youtube.com/watch?v=0Ig24JrzPek&t=17s

St. Croix River Education District. (2022, August 26-c). *A teacher's guide to the Be Good People SEL curriculum: Core edition, grades 4-5.* [Video]. YouTube. https://www.youtube.com/watch?v=KC-VkpK3oHA&t=1s

Starr, J. P. (2019). On leadership: Can we keep SEL on course? *Phi Delta Kappan*, *100*(8), 70–71. https://doi.org/10.1177/0031721719846894

Stoiber, K. C., & Gettinger, M. (2015). Multi-tiered systems of support and evidence-based practices. In *Handbook of response to intervention: The science and practice of multi-tiered systems of support* (pp. 121-141). Boston, MA: Springer US.

Sullivan, J., Wilton, L., & Apfelbaum, E. P. (2021). Adults delay conversations about race because they underestimate children's processing of race. *Journal of Experimental Psychology: General*, *150*(2), 395.https://doi.org/10.1037/xge0000851

Taylor, R. D., Oberle, E., Durlak, J. A., & Weissberg, R. P. (2017). Promoting positive youth development through school-based social and emotional learning interventions: A meta-analysis of follow-up effects. *Child development*, *88*(4), 1156-1171. https://doi-org.ezp3.lib.umn.edu/10.1111/cdev.12864

Theoharis, G. (2008). "At every turn": The resistance that principals face in their pursuit of equity and justice.. *Journal of School Leadership, 18*(3). https://doi-org.ezp3.lib.umn.edu/10.1177/105268460801800304

Thornhill, C. W., Castillo, L. G., Piña-Watson, B., Manzo, G., & Cano, M. À. (2021). Mental health among Latinx emerging adults: Examining the role of familial accusations of assimilation and ethnic identity. *Journal of Clinical Psychology, 78*(5), 892-912. https://doi.org/10.1002/jclp.23271

Tobin, J. (1995). The irony of self-expression. *American Journal of Education*, *103*(3), 233–258. https://www.jstor.org/stable/1085531

Tomczak, M., & Tomczak-Lukaszewska, E. (2014). The need to report effect size estimates revisited. An overview of some recommended measures of effect size. *Trends in Sport Sciences, 21*(1), 19-25. https://www.researchgate.net/publication/303919832_The_need_to_report_effect

_size_estimates_revisited_An_overview_of_some_recommended_measures_of_effect_size

Tyton Partners. (2021). *Finding your place in 2021: SEL takes center stage in K-12.* https://d1hzkn4d3dn6lg.cloudfront.net/production/uploads/2021/10/Tyton-Partners_Finding-Your-Place-2021_SEL-Takes-Center-Stage-in-K12.pdf

Ura, S. K., Castro-Olivo, S. M., & d'Abreu, A. (2020). Outcome measurement of school-based sel intervention follow-up studies. *Assessment for Effective Intervention*, *46*(1), 76–81. https://doi+.org/10.1177/1534508419862619

Van Oss (2023). *Be good people registration map.* Google Maps. https://www.google.com/maps/d/viewer?mid=1CnmrGvvH00FYDN6rstoiAOa2hdEU6D4&femb=1&ll=-3.81666561775622e-14%2C-116.60215090000003&z=1

Van Oss, N. (2024). *Be Good People registration map.* Google Maps. https://www.google.com/maps/d/u/0/viewer?mid=1CnmrGvvH00FYDN6rstoiAOa2hdEU6D4&femb=1

Vincze, L., Barnes, K., Somerville, M., Littlewood, R., Atkins, H., Rogany, A., & Williams, L. T. (2021). Cultural adaptation of health interventions including a nutrition component in Indigenous peoples: A systematic scoping review. *International Journal for Equity in Health*, *20*(1), 125. https://doi.org/10.1186/s12939-021-01462-x

von der Embse, N. P., Pendergast, L. L., Kilgus, S. P., & Eklund, K. R. (2016). Evaluating the applied use of a mental health screener: Structural validity of the Social, Academic, and Emotional Behavior Risk Screener. *Psychological Assessment, 28*(10), 1265–1275. https://doi-org.ezp3.lib.umn.edu/10.1037/pas0000253

Vygotsky, L. S. (2012). *Thought and Language, revised and expanded edition*. MIT Press. https://books.google.com/books?hl=en&lr=&id=rgr-

DwAAQBAJ&oi=fnd&pg=PR4&ots=acUaKJKJKg&sig=SRBJri32omozZCcwMegnfWU6dAw#v=onepage&q&f=false

Ward Schofield, J. (2019). The colorblind perspective in school: Causes and consequences. In Banks, J. A., & McGee, C. A. (Eds.), *Multicultural education: Issues and perspectives* (pp. 259-284). Wiley. https://books.google.com/books?hl=en&lr=&id=e1ITbOA2jhQC&oi=fnd&pg=PA259&dq=color+blind+perspective&ots=8Ak1VuO0Oa&sig=9pDJli5LSKHC4eb4r9sT2xCspPA#v=onepage&q=color%20blind%20perspective&f=false

Watson, K. R., Capp, G., Astor, R. A., Kelly, M. S., & Benbenishty, R. (2022). "We Need to Address the Trauma": School Social Workers' Views About Student and Staff Mental Health During COVID-19. *School mental health*, *14*(4), 902-917. https://doi.org/10.1007/s12310-022-09512-7

Wigelsworth, M., Mason, C., Verity, L., Qualter, P., & Humphrey, N. (2021). Making a case for core components: New frontiers in SEL theory, research, and practice. *School Psychology Review*, 1-14. https://doi.org/10.1080/2372966X.2021.2004863

Will, M. (2022, April). *What should culturally relevant teaching look like today? Gloria Ladson-Billings explains.* Education Week. https://www.edweek.org/leadership/what-should-culturally-relevant-teaching-look-like-today-gloria-ladson-billings-explains/2022/04

Xie, Y., & Greenman, E. (2011). The social context of assimilation: Testing implications of segmented assimilation theory. *Social Science Research, 40*(3), 965-884. https://doi.org/10.1016/j.ssresearch.2011.01.004

Yarin, S. (2022, October 7). *If Hispanics Hate the Term "Latinx", Why is it Still used?* Boston University. https://www.bu.edu/articles/2022/why-is-latinx-still-used-if-hispanics-hate-the-term/

Appendix A

Codebook

Key:

- MISSING: The information is not mentioned in the article at all
- NOT EXAMINED: The authors explicitly mention they did not analyze a variable
- NOT REPORTED: The authors had a measure, but did not report findings
- NOT REPORTED AS EXPECTED: The way the data were reported were not in the same unit of study as other variables (e.g. reports demographic for the entire school instead of for the group)

IMPORTANT: if multiple studies or interventions that are described, provide the following information only for the study with the adapted intervention

General Information			
Variable	Variable Definition	Codes	Notes
ID	Study identification number	Enter study identification number	
YEAR	Publication year	Enter publication year	
AUTH	Study author(s)	Enter (up to) first three last names and first initials of authors, and “et al.” for remaining authors (if applicable)	Example for one to three authors: Li, A., Miller., F Example for more than three authors: Li, A., Miller, F., Reinke, P. et al.

TITLE	Article title	Enter full article title	
RQ	Research questions	Enter research questions. Number the questions in a bulleted list if there is more than one question	NOTE: If the research question(s) is/ are only stated in purpose form (e.g. “the purpose of this study is to…”) then enter in the purpose of the study.
DESIGN	Study Design	A = Between group B = Within group C = other (describe) AND 0 = single-case design 1 = quasi-experimental 2 = randomized control trial 3 = qualitative 4 = other (describe)	Single-case design examples include multiple baseline, ABAB, alternating treatment, multiple treatment, etc
SEL Program Details			
PNAME	Program name	Enter name of SEL program implemented (if applicable)	
PTYPE	Type of SEL program used (ie. how the program is being used in the study)	0 = individual program 1 = small-group 2 = class-wide 3 = school-wide 4 = other (describe)	NOTE. if multiple formats are used, select “other” and describe

PLESS	Number of lessons/sessions (or activities if lesson/session numbers are not available)	Enter the number of lessons/sessions that were conducted. If the specific number of lessons/sessions are not provided, provide detail of how long sessions are or any other related information.	
PTIME	Duration of program	Enter the number of days, weeks, or months the SEL program is implemented	NOTE. If the duration of program is not described in days, weeks, or months, enter the duration of the program as described by the author(s)
CASEL	CASEL competency targeted by the SEL program	A = CASEL competency is explicitly targeted (ie. says targeting "self-management") B = CASEL competency is implicitly targeted C = other skills or competencies targeted by the adapted SEL program (describe) **AND** 0 = self-management 1 = self awareness 2 = responsible decision-making 3 = relationship skills 4 = social awareness 5 = multiple (list which ones)	NOTE: Code both A or B AND the numbers (e.g. A1) "Other" can be used to describe any other competencies that do not fall under a CASEL competency Provide examples if the CASEL competency is implicitly targeted. Program skills can fall under multiple competencies Definitions from CASEL: **Self-management:** The abilities to manage one's emotions, thoughts, and behaviors effectively in different situations and to achieve

			goals and aspirations. This includes the capacities to delay gratification, manage stress, and feel motivation and agency to accomplish personal and collective goals. **Self-awareness:** The abilities to understand one's own emotions, thoughts, and values and how they influence behavior across contexts. This includes capacities to recognize one's strengths and limitations with a well-grounded sense of confidence and purpose. **Responsible decision-making:** The abilities to make caring and constructive choices about personal behavior and social interactions across diverse situations. This includes the capacities to consider ethical standards and safety concerns, and to evaluate the benefits and consequences of various actions for personal, social, and collective well-being. **Relationship skills:** The abilities to establish and maintain healthy and supportive relationships and to

			effectively navigate settings with diverse individuals and groups. This includes the capacities to communicate clearly, listen actively, cooperate, work collaboratively to problem solve and negotiate conflict constructively, navigate settings with differing social and cultural demands and opportunities, provide leadership, and seek or offer help when needed. **Social awareness:** The abilities to understand the perspectives of and empathize with others, including those from diverse backgrounds, cultures, and contexts. This includes the capacities to feel compassion for others, understand broader historical and social norms for behavior in different settings, and recognize family, school, and community resources and supports.
Participant and Implementer Characteristics			
PP_T	Total number of participants in the study	Enter the number of participants in the study	NOTE: Only include total number of participants in the final analysis if there was attrition (applies to all

			participant and implementer characteristics) NOTE: If there is an experimental and control group, provide data for both separately and note which group the data comes from (applies to all participant and implementer characteristics) EXAMPLE: N = 47 (control); N = 46 (experimental)
PP_GRADE	Grade(s) of the participants	0 = PreK 1 = K-5 (ie. elementary) 2 = 6-8 (ie. middle/junior high) 3 = 9-12 (ie. high school) 4 = other (describe)	NOTE: if there are multiple grades (e.g. K-5 and 6-8; 4-6) enter "other code" and describe as is.
PP_RE	Race/ethnicity of the participants and their percentages	Enter race/ethnicity data as the authors reported. Provide the percentages for each racial/ethnic category presented (may need to calculate)	Example: Asian American (10%), Hispanic (35%)
PP_GEN	Gender or sex assigned at birth of the participants and their percentages	Enter gender or sex assigned at birth data as the authors reported. Provide the percentages for each gender or sex category presented (may need to calculate)	Example: male (40%); female (30%)

I_T	Total number of implementers in the study	Enter the number of implementers in the study	
I_R	Who is implementing the intervention (ie. what role do they hold at the school?)	0 = Teacher 1 = School counselor, psychologist, social worker, or nurse 2 = Education support staff 3 = Administrator 4 = Researcher 5 = Other (describe)	
I_EXP_R	Number of years of experience in their role	0 = Less than 5 1 = 5-10 years 2 = 10+ years 3 = other (describe)	NOTE: “role” is defined as by the I_R code
I_EXP_SEL	Number of years of experience with the implemented SEL program	0 = No experience 1 = 1 day -5 years 2 = 5-10 years 3 = 10+ years 4 = other (describe)	*if training is provided, code as “other” and provide length of training
I_RE	Race/ethnicity of the implementer(s)	Enter race/ethnicity data as the authors reported. Provide the percentages for each racial/ethnic category presented (may need to calculate)	Example: Asian American (10%), Hispanic (35%)
I_GEN	Gender of the implementer(s)	Enter gender or sex assigned at birth data as the authors reported. Provide the percentages for each gender or sex category presented (may need to calculate)	Example: male (40%); female (30%)

Method Details			
SETT	Setting of the school(s) and/or district(s)	If available, enter the number of schools and/or districts in the study (e.g. N = 3 schools in N = 2 districts OR N = 1 school OR N = 5 districts). If available, list the name of the city, state, or region **AND** code for the following: If available, was each school/district: 0 = rural 1 = suburban 2 = urban 3 = other (describe)	NOTE: if there are multiple settings, code as "other" and describe Example: N = 30 schools,, Minnesota, 2 (ie. there were 30 schools in minnesota that were urban)
ATYPE	Adaptation type to the SEL program	**1 = Surface Structure Feature (Presentation)** A = Bilingual materials B = Bicultural materials C = Bilingual implementer D = Bicultural implementer E = Translation/back-translation of materials F = Inclusion of ethnic lifestyle elements (e.g. foods, music) G = Including cultural brokers H = Delivering SEL program in group settings (note. only if the intervention is implemented individually AND reasoning is provided for adapting the intervention into a group setting) I = conducting the SEL program in safe locations familiar to the participants J = Hands on activities	NOTE: Code with 1 OR 2 AND THEN the letter in the corresponding section (Example: 1A) NOTE. Multiple adaptations in a study can be noted. If there are multiple adaptations, list them in a list format (Example: - 1A - bilingual materials - 1D - implementer was bicultural - 2N - included a cultural value into the activities) NOTE: Please list ALL adaptations

		K = Incorporating culturally familiar formats and activities (e.g. novela format, culture specific dances) L = inclusion of same-race/ethnicity role model M = other (describe) —----- **2 = Deep Structure Feature (Content)** N = incorporating cultural values in SEL program design O = incorporating cultural values in SEL program implementation P = involving family in SEL program (ONLY if this is not part of the SEL program originally AND reasoning is provided for including family in the SEL program) Q = adjusting materials to literacy level of participant R = using social support and/or networks S = other (describe)	made to the program that are not typically part of the program (ie. if the directions call for implementers to adapt the script to sound more like the implementer, DO NOT include this as an adaption) NOTE: Give brief description of what the adaptation was when coding (Example: 1A = had worksheets in Spanish and English) NOTE: **"Surface structure components** include intervention materials, messages, channels, settings, and recruitment strategies that match the characteristics of a priority population…**Deep structure components** are cultural, social, historical, environmental, and psychological factors influencing health behaviors." (Mier et al., 2010)
APURP	Purpose for the adaptation to the SEL program	Enter the author(s) stated purpose for creating the adaptation(s)	
SO_MEASURE	Measure(s) of the dependent variable(s) that were used related to	Enter the full name of all the measures that were used to measure student outcomes and any applicable acronyms	NOTE. Measures should be related to student outcomes (ie. social, emotional, behavioral, academic

	student outcomes		outcomes) NOTE: If no measure name is given, provide a short, author description, if it is author developed, and/or source of scale (e.g. OUTCOME MEASURED, (from Li (2023))
SO_DV	Dependent variable(s), that are related to student outcomes associated with the cultural adaptation to the SEL program	Enter the dependent variable(s) of the study related to student outcomes	Example: Student GPA NOTE. If there are several dependent variables, put an * by the primary dependent variable of investigation, if authors noted this (Example: grades*, confidence level, relatability)
SO_ES	Student outcome effect sizes associated with the cultural adaptations	Enter the Cohen's D effect size (or other effect sizes reported by author's) of the cultural adaptation on the measured student outcome and their context (ie. What DV is the effect size in reference to? What does the effect size reflect or demonstrate? Was the effect size a difference between pre and post measurements? Was the effect size a difference between groups? Provide enough information to understand what analyses were completed). If Cohen's D effect sizes are not available, report and contextualize any other values or qualitative information related to student outcomes and contextualize it.	NOTE. Provide all reported data related to the culturally adapted intervention NOTE. The effect size(s) may need to be calculated. If the effect size needs to be calculated, use this website and only calculate pre-test and post-test differences. It is not necessary to calculate for any follow-up conditions. The calculator can only be used for between group differences. For within-group differences, just report and contextualize any other

			values or qualitative values related to student outcomes.
SV_YN	Social validity was measured	1 = yes 2 = no 3 = other (describe)	NOTE. if the answer is “no” then enter “N/A” for the rest of the SV measures NOTE. Social validity measures are related to the importance and accessibility of the goals, procedures, and outcomes (Foster & Mash, 1999) of the SEL program
SV_MEA_RES	Measures of social validity and their respondents	Enter the name or type of measures used to measure social validity and who responded to the measure in the following format: Measure (respondent/s)	Example: Interview (student, teacher); Rating Scale (student)
SV_DV	Dependent variables of social validity measures	Enter what the dependent variables of the social validity measures were in the following format: Measure (dependent variable)	Example: Interview (Feasibility for teacher implementation)
SV_OUT	Outcomes of social validity measures	Enter the outcomes of the social validity measures in the following format: Measure (outcome)	Example: Student Interview (Authors found the cultural adaptation improved student confidence)
TI	Treatment integrity	Enter the percentage of treatment integrity reported by the author(s) **AND** indicate how treatment integrity was measured and who measured the treatment integrity	Example: 100%; measured by a 3-item fidelity checklist completed by teacher NOTE: The percentage may need to

			be calculated and multiple forms of measures may be used, list all forms of measures NOTE: If only descriptives provided and no percentages, write "NO MEASURE, QUALITATIVE ONLY" and provide a general description of the fidelity (e.g. "NO MEASURE, QUALITATIVE ONLY: in general, most teachers implemented with fidelity)
Quality Appraisal			
THEORY	Theoretical or conceptual underpinning to the research	0 = No mention at all 1 = General reference to broad theories or concepts that frame the study. e.g. key concepts were identified in the introduction section. 2 = Identification of specific theories or concepts that frame the study and how these informed the work undertaken. e.g. key concepts were identified in the introduction section and applied to the study. 3 = Explicit discussion of the theories or concepts that inform the study, with application of the theory or concept evident through the design, materials and outcomes explored. e.g.	

		key concepts were identified in the introduction section and the application apparent in each element of the study design.	
STATE_RA	Statement of Research Aim(s)	0 = No mention at all 1 = Reference to what the sought to achieve embedded within the report but no explicit aims statement. 2 = Aims statement made but may only appear in the abstract or be lacking detail. 3 = Explicit and detailed statement of aim/s in the main body of report	
DESCRIP_POP	Clear description of research setting and target population	0 = No mention at all 1 = General description of research area but not of the specific research environment e.g. 'in primary care. 2 = Description of research setting is made but is lacking detail e.g. 'in primary care practices in region [x]'. 3 = Specific description of the research setting and target population of study e.g. 'nurses and doctors from GP practices in [x] part of [x] city in [x] country.'	
SD_APPR	The study design is	0 = No research aim/s stated or the design is	

O	appropriate to address the stated research aim(s)	entirely unsuitable e.g. a Y/N item survey for a study seeking to undertake exploratory work of lived experiences. 1 = The study design can only address some aspects of the stated research aim/s e.g. use of focus groups to capture data regarding the frequency and experience of a disease 2 = The study design can address the stated research aim/s but there is a more suitable alternative that could have been used or used in addition e.g. addition of a qualitative or quantitative component could strengthen the design. 3 = The study design selected appears to be the most suitable approach to attempt to answer the stated research aim/s.	
SAMP_APP	Appropriate sampling to address the research aim/s	0 = No mention of the sampling approach. 1 = Evidence of consideration of the sample required e.g. the sample characteristics are described and appear appropriate to address the research aim/s. 2 = Evidence of consideration of sample required to address the aim. e.g. the sample characteristics are described with reference to the aim/s.	

		3 = Detailed evidence of consideration of the sample required to address the research aim/s. e.g. sample size calculation or discussion of an iterative sampling process with reference to the research aims or the case selected for study	
DT_CHOICE	Rationale for choice of data collection tool/s	0 = No mention of rationale for data collection tool used. 1 = Very limited explanation for choice of data collection tool/s. e.g. based on availability of tool. 2 = Basic explanation of rationale for choice of data collection tool/s. e.g. based on use in a prior similar study 3 = Detailed explanation of rationale for choice of data collection tool/s. e.g. relevance to the study aim/s, codesigned with the target population or assessments of tool quality.	
DT_APP	The format and content of data collection tool is appropriate to address the stated research aim/s	0 = No research aim/s stated and/or data collection tool not detailed. 1 = Structure and/or content of tool/s suitable to address some aspects of the research aim/s or to address the aim/s superficially e.g. single item response that is very general or an open-response item to capture content which requires probing.	

		2 = Structure and/or content of tool/s allow for data to be gathered broadly addressing the stated aim/s but could benefit from refinement. e.g. the framing of survey or interview questions are too broad or focused to one element of the research aim/s. 3 = Structure and content of tool/s allow for detailed data to be gathered around all relevant issues required to address the stated research aim/s.	
DATA_PROC	Description of data collection procedure	0 = No mention of the data collection procedure. 1 = Basic and brief outline of data collection procedure e.g. 'using a questionnaire distributed to staff' 2 = States each stage of data collection procedure but with limited detail or states some stages in detail but omits others e.g. the recruitment process is mentioned but lacks important details 3 = Detailed description of each stage of the data collection procedure, including when, where and how data was gathered such that the procedure could be replicated.	
RECRUIT	Recruitment data	0 = No mention of recruitment data.	

	provided	1 = Minimal and basic recruitment data e.g. number of people invited who agreed to take part 2 = Some recruitment data but not a complete account e.g. number of people who were invited and agreed. 3 = Complete data allowing for full picture of recruitment outcomes e.g. number of people approached, recruited, and who completed with attrition data explained where relevant.	
AMETH_ SELC	Justification for analytic method selected	0 = No mention of the rationale for the analytic method chosen 1 = Very limited justification for choice of analytic method selected. e.g. previous use by the research team. 2 = Basic justification for choice of analytic method selected e.g. method used in prior similar research. 3 = Detailed justification for choice of analytic method selected e.g. relevance to the study aim/s or comment around of the strengths of the method selected.	
AMETH_ APP	The method of analysis was appropriate to	0 = No mention at all.	

	answer the research aim/s	1 = Method of analysis can only address the research aim/s basically or broadly 2 = Method of analysis can address the research aim/s but there is a more suitable alternative that could have been used or used in addition to offer a stronger analysis. 3 = Method of analysis selected is the most suitable approach to attempt answer the research aim/s in detail e.g. for qualitative interpretative phenomenological analysis might be considered preferable for experiences vs. content analysis to elicit frequency of occurrence of events.	
CONS	Evidence that the research stakeholders have been considered in research design or conduct.	0 = No mention at all. 1 = Consideration of some of the research stakeholders e.g. use of pilot study with target sample but no stakeholder involvement in planning stages of study design. 2 = Evidence of stakeholder input informing the research. e.g. use of pilot study with feedback influencing the study design/conduct or reference to a project reference group established to guide the research. 3 = Substantial consultation with stakeholders identifiable in planning of study design and in	

		preliminary work e.g. consultation in the conceptualisation of the research, a project advisory group or evidence of stakeholder input informing the work.	
SL_DISCUSS	Strengths and limitations critically discussed	0 = No mention at all. 1 = Very limited mention of strengths and limitations with omissions of many key issues. e.g. one or two strengths/limitations mentioned with limited detail. 2 = Discussion of some of the key strengths and weaknesses of the study but not complete. e.g. several strengths/limitations explored but with notable omissions or lack of depth of explanation. 3 = Thorough discussion of strengths and limitations of all aspects of study including design, methods, data collection tools, sample & analytic approach.	
QUADS_TOTAL	Total Score on the QUADS	Enter the total number of points scored on the QUADS tool	

Appendix B

Quality Assessment for Diverse Studies (QuADS): User Guide v1.0

Step 1: Establish whether QuADS is relevant for your review

The QuADS tool is developed for use in literature reviews that include multi-method or mixed-methods research. The tool is applied to the studies that are selected for inclusion in the review based on the eligibility criteria. Whilst the tool can be applied to reviews that have a group of quantitative and a group of qualitative studies within, it is also possible to assess each group separately and then conduct a synthesised quality assessment as an alternate approach. It is important to consider the relevance of each of the available tools to your group of papers to determine which approach will allow you to optimally assess the body of work. The QuADS tool can also be used with reviews of multi- or mixed methods work and some examples are given as to how the criteria could be applied in the context of a review article.

Step 2: Agree the way in which your team will apply the QuADS criteria

Once the review team have finalised the studies to be included in the review, the team should first discuss the way in which the tool will be applied to the resultant studies. The QuADS tool requires reviewers to apply their own research knowledge and judgement in making decisions around the appropriate scoring for each study. An optimal process occurs when the reviewer team collectively discuss the QuADS criteria and how they might apply these in the context of the body of work of focus. The facilitates reviewers to develop a shared understanding of the items in the context of their review.

Step 3: Independently apply the QuADS criteria to a small group of included studies

Exploring the quality of research reporting and conduct requires an iterative process to enable the review team to consolidate a shared understanding of what high quality reporting looks like in their given field or research topic. The team members should independently try to apply the QuADS tool to the same 3-5 articles. The scores are applied on a four-point scale to enable researchers to distinguish the degree to which a criteria is met.

Step 4: Instigate reviewer discussion

Once the criteria have been applied to a small number of studies, the review team can then meet to discuss their scoring and experiences. Through this process, the team will finalise a shared understanding of how they will apply these criteria and discuss any inconsistencies or challenges in application of the criteria to particular studies or the group of studies.

Step 5: Independently apply the QuADS criteria to a further group of included studies or the remainder

Once agreement around the application of the tool to the work has been reached through steps 3 and 4, each reviewer can then apply the criteria independently to all studies in the review, or this may be a subset in the context of reviews of a high volume of papers. In the context of substantial disagreement or difficulty in applying the criteria to any given study or body of work, it may be necessary to repeat steps 3 and 4 to consolidate an understanding of how to apply the tool to the given body of work.

Step 6: Review the scores obtained across the body of work against each criteria

Once all papers have been reviewed by all reviewers, the team once again meet to discuss their scoring, any discrepancies and those discrepancies that may be resolved. The reviewers will finalise their scores and may then wish to undertake an inter-rater reliability analysis with these.

Step 7: Using the resulting data

The resulting data from the QuADS tool provides researchers an opportunity to consider the reporting of information across the body of work related to each criterion. Criteria are not weighted but some may be considered by the research team to be particularly important to their research field and these findings can be discussed narratively when writing up findings regarding the study quality assessment.

There is also no cut-off score for a study to be considered high or low quality; any cut-off would be arbitrary and not appropriate when using this tool. Through reviewing scores for each single criterion across the body of work rather than the overall scores, researchers are advised to discuss the quality assessment findings narratively and consider areas in which reporting is comprehensive or less so and why this may be.

Appendix C

QUADS Quality Tool Table

QuADS Criteria	0	1	2	3
1. **Theoretical or conceptual underpinning to the research**	No mention at all.	General reference to broad theories or concepts that frame the study. e.g. key concepts were identified in the introduction section.	Identification of specific theories or concepts that frame the study and how these informed the work undertaken. e.g. key concepts were identified in the introduction section and applied to the study.	Explicit discussion of the theories or concepts that inform the study, with application of the theory or concept evident through the design, materials and outcomes explored. e.g. key concepts were identified in the introduction section and the application apparent in each element of the study design.
2. Statement of research aim/s	No mention at all.	Reference to what the sought to achieve embedded within the report but no explicit aims statement.	Aims statement made but may only appear in the abstract or be lacking detail.	Explicit and detailed statement of aim/s in the main body of report.
3. Clear description of research setting and target population	No mention at all.	General description of research area but not of the specific research environment e.g. 'in primary care.'	Description of research setting is made but is lacking detail e.g. 'in primary care practices in region [x]'.	Specific description of the research setting and target population of study e.g. 'nurses and doctors from GP practices in [x] part of [x] city in [x] country.'
4. The study design is appropriate to address the stated research aim/s	No research aim/s stated or the design is entirely unsuitable e.g. a Y/N item survey for a study seeking to undertake exploratory work of lived experiences. .	The study design can only address some aspects of the stated research aim/s e.g. use of focus groups to capture data regarding the frequency and experience of a disease.	The study design can address the stated research aim/s but there is a more suitable alternative that could have been used or used in addition e.g. addition of a qualitative or	The study design selected appears to be the most suitable approach to attempt to answer the stated research aim/s.

			quantitative component could strengthen the design.	
5. Appropriate sampling to address the research aim/s	No mention of the sampling approach.	Evidence of consideration of the sample required e.g. the sample characteristics are described and appear appropriate to address the research aim/s.	Evidence of consideration of sample required to address the aim. e.g. the sample characteristics are described with reference to the aim/s.	Detailed evidence of consideration of the sample required to address the research aim/s. e.g. sample size calculation or discussion of an iterative sampling process with reference to the research aims or the case selected for study.
6. Rationale for choice of data collection tool/s	No mention of rationale for data collection tool used.	Very limited explanation for choice of data collection tool/s. e.g. based on availability of tool.	Basic explanation of rationale for choice of data collection tool/s. e.g. based on use in a prior similar study.	Detailed explanation of rationale for choice of data collection tool/s. e.g. relevance to the study aim/s, co-designed with the target population or assessments of tool quality.
7. The format and content of data collection tool is appropriate to address the stated research aim/s	No research aim/s stated and/or data collection tool not detailed.	Structure and/or content of tool/s suitable to address some aspects of the research aim/s or to address the aim/s superficially e.g. single item response that is very general or an open-response item to capture content which requires probing.	Structure and/or content of tool/s allow for data to be gathered broadly addressing the stated aim/s but could benefit from refinement. e.g. the framing of survey or interview questions are too broad or focused to one element of the research aim/s.	Structure and content of tool/s allow for detailed data to be gathered around all relevant issues required to address the stated research aim/s.
8. Description of data collection procedure	No mention of the data collection procedure.	Basic and brief outline of data collection procedure e.g. 'using a questionnaire distributed to staff'.	States each stage of data collection procedure but with limited detail or states some stages in detail but	Detailed description of each stage of the data collection procedure, including when, where and how

			omits others e.g. the recruitment process is mentioned but lacks important details.	data was gathered such that the procedure could be replicated.
9. Recruitment data provided	No mention of recruitment data.	Minimal and basic recruitment data e.g. number of people invited who agreed to take part.	Some recruitment data but not a complete account e.g. number of people who were invited and agreed.	Complete data allowing for full picture of recruitment outcomes e.g. number of people approached, recruited, and who completed with attrition data explained where relevant.
10. Justification for analytic method selected	No mention of the rationale for the analytic method chosen.	Very limited justification for choice of analytic method selected. e.g. previous use by the research team.	Basic justification for choice of analytic method selected e.g. method used in prior similar research.	Detailed justification for choice of analytic method selected e.g. relevance to the study aim/s or comment around of the strengths of the method selected.
11. The method of analysis was appropriate to answer the research aim/s	No mention at all.	Method of analysis can only address the research aim/s basically or broadly.	Method of analysis can address the research aim/s but there is a more suitable alternative that could have been used or used in addition to offer a stronger analysis.	Method of analysis selected is the most suitable approach to attempt answer the research aim/s in detail e.g. for qualitative interpretative phenomenological analysis might be considered preferable for experiences vs. content analysis to elicit frequency of occurrence of events.
12. Evidence that the research stakeholders have been considered in research design or conduct.	No mention at all.	Consideration of some the research stakeholders e.g. use of pilot study with target sample but no	Evidence of stakeholder input informing the research. e.g. use of pilot study with feedback influencing the study	Substantial consultation with stakeholders identifiable in planning of study design and in preliminary work e.g. consultation in the

		stakeholder involvement in planning stages of study design.	design/conduct or reference to a project reference group established to guide the research.	conceptualisation of the research, a project advisory group or evidence of stakeholder input informing the work.
13. Strengths and limitations critically discussed	No mention at all.	Very limited mention of strengths and limitations with omissions of many key issues. e.g. one or two strengths/limitations mentioned with limited detail.	Discussion of some of the key strengths and weaknesses of the study but not complete. e.g. several strengths/limitations explored but with notable omissions or lack of depth of explanation.	Thorough discussion of strengths and limitations of all aspects of study including design, methods, data collection tools, sample & analytic approach.

Appendix D

Complete list of Kurtz (2020) Adaptations

This complete list was compiled from Appendix B of the Kurtz (2020) dissertation (pg. 106-124).
The adaptations below were grouped by their adaptation type.

Surface level: Translation/Back-translation of materials:

- Family bulletin translated into Spanish to send home with students
- The Way I Feel (Spanish version) to be shared with students

Surface level: Inclusion of ethnic lifestyle elements:

- Ask students to share feeling words in the language they use at home
- Ask students to share synonyms for angry in the language they use at home
- Ask students to share synonyms for happy in the language they use at home
- Ask students to share synonyms for worried in the language they use at home

Surface level: Inclusion of same-race/ethnicity role model:

- Choose an alternative book for this lesson where the main character is a child of color: Marvelous Me by Lisa Bullard
- Choose an alternative book for this lesson where the main character is a child of color: Those Shoes by Maribeth Boelts
- Choose an alternative book for this lesson where the main character is a child of color: Listening With My Heart by Gabi Garcia

Surface level: Other

- Send email with common language used during group to small group teachers following group
- Incorporate increased opportunities for students to respond (e.g., use clear sheet protectors with handouts in them and dry erase markers for students to respond to questions posed by the interventionist)
- Include additional mindfulness activity and coping skill instruction (e.g., breathing, squeezing clay)
- Include a mindfulness break at the beginning and end of the group
- Offer additional coping tools (e.g., elevator breathing, squeezing clay)
- Incorporate an activity focused on feelings identification, active listening, empathy, and turn-taking where students take turns choosing a colored Goldfish and sharing whether the corresponding feeling on the key provided makes them feel “good” or “not so good” and what might may cause that feeling
- Use teacher feedback to refer to “expected” and “unexpected” ways of expressing feelings instead of “okay” and “not okay” ways of expressing feelings
- Include turn-taking activity where students try to throw a bean bag onto one of multiple scenarios (e.g., someone yells at you) on the floor and they are asked to share an expected way of coping with the feeling

- Include extension activity of "Jumping Bean Bags" but instead of "A, B, C, or D" use the visuals for the six basic feelings faces on the floor for students to throw their bean bags on in response to the scenarios posed
- Include extension activity: "Body Tracing" with individual body maps for each student
- Incorporate an activity focused on practicing sunshine thoughts, active listening, empathy, and turn-taking where students take turns choosing a colored Goldfish and practicing sharing a sunshine thought or cloudy thought
- Use individual body maps to identify how feelings are experienced in the body
- Incorporate a listening activity where students have to repeat back what a peer said in response to question posed by the interventionist
- Instead of using the terminology "nice voice" and a "not so nice voice", "caring voice" was used
- Incorporate the "Size of the Problem" visual from The Zones of Regulation curriculum (Kuypers, 2011) as another coping tool students can use to cope with feelings and solve problems
- Include an activity where students take turns being the "Tiny Problem" superhero and wearing a cape and crown to turn a problem into a "tiny" or "little" with scenarios teachers identified prior to group as being triggering for students
- Build in opportunities for students to experience mastery of a problem that may seem large or scary
- Adapt "Strong Start Toolbox" extension activity to be a coping toolbox and have students include visuals of sunshine thoughts, StopCount-In-Out, the size of problems, big deal vs. little deal, and additional coping tools
- Emphasize students can utilize these coping toolboxes when they feel out of control and/or are thinking about "hard things"
- Introduce the idea of termination/group ending and emphasize it has nothing to do with the students and that it is a positive goodbye
- Choose an alternative book for this lesson: Emotes! The Emotes BIG Book of Feelings by Matt Casper
- Include certificates to highlight positive nature of ending of group
- Include a body mapping activity where each student has an individual body map they can use to identify how they experience certain emotions when the interventionist poses different scenarios
- Teach the trauma response of "flight, fight, or freeze"
- Incorporate an activity focused on feelings identification, active listening, empathy, and turn-taking where students take turns choosing a colored Goldfish and sharing whether the corresponding feeling on the key provided makes them feel "comfortable" or "uncomfortable" and what might may cause that feeling
- Include turn-taking activity where students try to throw a bean bag onto one of multiple scenarios (e.g., someone yells at you) on the floor and they are asked to share a helpful way of coping with the feeling
- Incorporate the visual of an outline of a pair of shoes for stepping into someone else's shoes to support the development of students' understanding of empathy as a concept

- Include extension activity of "Jumping Bean Bags" from Lesson 4 of Strong Start but instead of "A, B, C, or D" use the visuals for the six basic feelings faces on the floor for students to throw their bean bags on in response to the scenarios posed
- Include a body mapping activity where each student has an individual body map they can use to identify how they experience anger in the body when the interventionist poses different scenarios
- Incorporate an activity focused on feelings identification, active listening, empathy, and coping skills identification where students take turns choosing a colored Goldfish and sharing relevant anger management skills as practiced during the lesson
- Incorporate Fortune Teller activity from Lesson 6 of Strong Start with sunshine thoughts identification as the focus of the fortune tellers
- Include an activity where students begin designing their own coping "toolboxes" where they add visuals representing the coping skills we have discussed during lessons
- Ensure common language from Second Step is utilized for conflict resolution steps
- Emphasize the role of technology in students' daily lives
- Discuss the impact of trauma on sleep and relaxation strategies that can be used
- Incorporate the extension activity of "Worry Stones" from Lesson 7 of Strong Start but alter the activity to involve the students designing stones related to make positive living choices
- Revisit the idea of termination
- Incorporate Bingo from Strong Start as a way to review concepts taught during lessons
- hoose an alternative book for this lesson: Help Your Dragon Deal With Anxiety by Steve Herman

Deep level: Incorporating cultural values in SEL program design:

- Asked for community feedback on adaptations*
- Use culturally relevant examples for having multiple feelings at once (e.g., when someone has to leave Boston)
- Choose Yo! Yes? as the book for this lesson
- Emphasize that individuals, families, and communities can have different ways of coping with emotions
- Emphasize that families have different ways of coping with anger
- Discuss the link between trauma and anger, and how trauma can make us feel like many situations/experiences are "big deals" despite them being "little deals"

Deep level: Incorporating cultural values in SEL program implementation

- Continue to incorporate culturally relevant examples for experiencing worry (e.g., when someone has to leave Boston, people yelling, scary noises in the neighborhood)
- Emphasize that families have different ways of coping with worries
- Emphasize that families have different ways of showing friendship and communicating, ask students to share different ways that families show they care

- Emphasize that communication and showing feelings often is different depending on where you are (home, school, community)
- se culturally relevant examples for triggers to emotions (e.g., when someone has to leave Boston or you have not seen them for awhile, loud noises in the neighborhood, the grown-ups you live with have a fight, you keep thinking about something that has happened in the past)
- Emphasize that communication and showing feelings often is different depending on who you are with and where you are (home, school, community)
- Have students share ways they have noticed how their family members and people in the community show empathy/support for others
- Continue to incorporate culturally relevant examples that may trigger anger (e.g., when someone has to leave Boston or you have not seen them for awhile, loud noises in the neighborhood, the grown-ups you live with have a fight, you keep thinking about something that has happened in the past, your parents are not able to read enough English to help you with your homework)
- Continue to incorporate culturally relevant examples that may trigger stress
- Discuss how individuals and families can have different ways of listening and communicating including during a conflict, how the context of where and with who you are experiencing a conflict with impacts how you respond to the conflict
- Provide an example of setting a goal that is relevant to the students' contexts, discuss goals they feel their families have for them

Deep level: Adjusting materials to literacy level of participant

- Incorporate the "additional activity" of the use of mirrors during emotion identification to support English Language Learners and students on the Autism Spectrum who may benefit from more visuals
- Utilize "sunshine thoughts" as a coping to use in response to thinking traps to decrease the verbal demand for English Language Learners
- Incorporate extension activity "Mirror Mirror" to support English Language Learners who may require more visuals
- Utilize "sunshine thoughts" instead of the ABCs of positive thinking to decrease the verbal demand for English Language Learners

Deep level: Using social support and/or networks:

- Emphasize safety in coping with anger and that adults are available at school to help children cope with anger, how to ask adults for help across settings
- Discuss ways to seek help with coping with trauma, depression, or anxiety and what to do if you tell an adult who seems like they cannot help you cope

Deep level: Other

- Include psychoeducation about how experiencing trauma can make it difficult to identify feelings when all feelings can feel like angry or scared
- Ask students to share physical clues related to how people at home show different emotions

- Include psychoeducation about how experiencing trauma can make it difficult to identify physical clues of emotions in others
- Teach link between experiencing trauma ("hard things") and feelings (e.g., how it might feel to have a family member have to move away or leave for another reason, you have a nightmare, you hear scary noises in the neighborhood)
- Teach link between experiencing trauma ("hard things"), the brain, and feelings (e.g., how it might feel to have a family member have to move away or leave for another reason, you have a nightmare, you hear scary noises in the neighborhood)
- Address how trauma can make us get stuck in thinking "cloudy thoughts" but that sunshine thoughts can help us become unstuck
- Include a current video clip from YouTube about coping with anger
- Teach the link between trauma and difficulty recognizing others' emotions and showing support
- Discuss positive choices that can be made in an urban environment where places to exercise may be more limited

*This was not an adaptation listed in Kurtz (2020) Appendix B, however, it was an adaptation to the program having community folk review and revise the research-proposed adaptations.

Appendix E

Teacher Demographics Questionnaire

Please answer the following questions to the best of your ability.

1. What is your teacher study ID?
    a. [Redacted for anonymity]
    b. [Redacted for anonymity]
    c. [Redacted for anonymity]
    d. [Redacted for anonymity]
    e. [Redacted for anonymity]
    f. [Redacted for anonymity]
    g. [Redacted for anonymity]
2. What is your gender identity
    a. Woman
    b. Man
    c. Non-binary
    d. Prefer not to disclose
    e. Other
3. What is your race/ethnicity
    a. White
    b. Black or African American
    c. Indigenous (ie. American Indian) or Alaska Native
    d. Asian
    e. Native Hawaiian and Other Pacific Islander
    f. Two or more races
    g. Prefer not to disclose
4. What is your highest degree earned?
    a. High school/GED
    b. Bachelor's
    c. Master's
    d. Doctorate
    e. Prefer not to disclose
    f. Other
5. Are you currently pursuing further education?
    a. Yes
    b. No
    c. Prefer not to disclose
    d. Other
6. If "yes" to the question "are you currently pursuing further education?," what degrees are you pursuing and in what subject area?
7. What grade do you currently teach?
    a. 1st

   b. 2nd
   c. 3rd
   d. 4th
8. Including this year, how many years have you taught your **current grade?** (Write this in digits)
9. Including this year, how many years have you been teaching **in total**? (Write this in digits)
10. How many students are in your classroom this year? (Write this in digits)
11. Including this year, how many years have you taught **the Be Good People curriculum?** (Write this in digits)
12. Did you support in the construction of the Be Good People curriculum? (Support = provided ideas, feedback, joined meetings when creating the curriculum, etc)
    a. Yes
    b. No
    c. Other

Appendix F

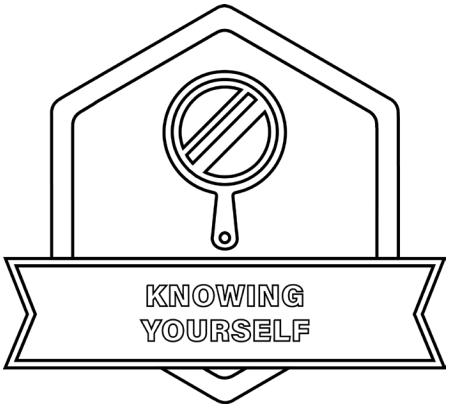
Weekly Teacher BGP Fidelity Log

Please fill this out every Friday. This survey will ask you weekly about the lessons and extension activities that you implemented. **Your individual data will not be available to administration.**

1. What is your teacher study ID?
    a. [Redacted for anonymity]
    b. [Redacted for anonymity]
    c. [Redacted for anonymity]
    d. [Redacted for anonymity]
    e. [Redacted for anonymity]
    f. [Redacted for anonymity]
    g. [Redacted for anonymity]
2. Which week is the following data for?
    a. Teacher selected the week from a drop down menu
3. Was there any change in your schedule this week? Examples include lessons not implemented at a typical time/day, drills interrupting lessons, teacher out for illness, etc. (write "NO" if this was a typical week)
4. How many LESSONS of BGP did you implement from beginning to end this week? (Enter as a number)
5. Which LESSON(s) did you implement? Please write the full name of the lesson.
6. How many EXTENSION ACTIVITIES did you implement this week? (Enter as a number)
7. Please answer the following questions regarding the **LESSON** you taught this week to the best of your knowledge (0 Not at all | 1 = Partially | 2 = Fully)
    a. To what extent were you able to lead through all of the slides?
    b. To what extent were you able to complete activities on all slides?
    c. To what extent did you have all students fill in learning logs at the end of the weekly lesson?
8. If you did not implement any Be Good People programming this week, check this box and submit the survey.
    a. Teacher checked the box if no lesson was implemented during the week

Appendix G

Weekly Student BGP Learning Log



Date (Month/Day)		Date (Month/Day)		Date (Month/Day)	

Appendix H

Researcher BGP Observation Form

Please note: The original sections of the instructional coaching observation form that BGP created were labeled as "supportive climate," "student engagement," "fidelity," and "delivery." However, to specify what each of the items more closely captured, the researchers labeled the sections as "supportive climate," "student engagement," "adherence," and "quality of instruction," respectively.

General Information

1. Who is observing
    a. Observer 1
    b. Observer 2
2. Date of observation
    a. Researcher selected the date of the observation
3. Time of observation
    a. Researcher selected the time of the observation
4. Teacher you are observing
    a. Teacher A
    b. Teacher B
    c. Teacher C
    d. Teacher D
    e. Teacher E
    f. Teacher F
    g. Teacher G
5. Which are you observing?
    a. Lesson
    b. Extension activity
    c. Flexible activity
6. Title of lesson/activity
7. What skill does this lesson cover?
    a. Making good choices
    b. Getting along with others
    c. Thinking of others
    d. Knowing yourself
    e. Managing yourself

Supportive Climate Section

0 = Not Evident | 1 = Partially Evident | 2 = Clearly Evident

1. Instructor primes the group before discussion by reviewing positively-stated expectations. Expectations are clearly visible and referenced again if needed.
2. Expectations are clearly visible and referenced again if needed.

3. The majority of students demonstrates active listening when conversing with their peers.
4. The majority of students appear comfortable sharing.
5. Notes on supportive climate

Student Engagement Section

0 = Not Evident | 1 = Partially Evident | 2 = Clearly Evident

1. The majority of students appear attentive to the instruction.
2. A variety of students participate and respond to the instructor's questions.
3. Notes on student engagement

Adherence Section

0 = Not Evident | 1 = Partially Evident | 2 = Clearly Evident

1. Instructor lead through all of the slides
2. Instructor was able to complete activities on all slides
3. Instructor had all students fill in learning logs at the end of the weekly lesson
4. Notes on adherence

Quality of instruction section

0 = Not Evident | 1 = Partially Evident | 2 = Clearly Evident

1. Instructor appears enthusiastic and interested in the content.
2. Instructor authentically and naturally elaborates on the content, adding additional information or personal anecdotes.
3. Instructor appears comfortable, the pacing is appropriate, and transitions are smooth.
4. Instructor pacing is appropriate.
5. Instructor transitions between slides and activities are smooth.
6. Instructor prompts the students to make connections between the content and their personal experiences (e.g., asking if they've experienced something similar to what's shared in a video).
7. Instructor invites critical thinking (e.g., asking students if they agree or disagree with the content on a slide) and responds non-judgmentally to students.
8. Instructor responds non-judgmentally to students.
9. Instructor uses a variety of discussion techniques (e.g. whole-group, think-pair-share, ink-pair-share). **NOTE:** Score 0 for 1 technique, 1 for two techniques, and 2 for more than two techniques.
10. Instructor demonstrates active listening skills (e.g., eye contact, restatement).
11. Notes on quality of instruction:

—

1. Any other important notes?

Appendix I

Teacher Interview Script

Begin recording by saying "DATE / TIME / TEACHER ID"

Hello, thank you for meeting with me. I appreciate the time you are taking to discuss your perceptions of the Be Good People program with me. Although I have given you the interview questions in advance, I may ask follow up questions to your answers that you provide today that were not provided in advance of this interview. If you do not want to answer any of the questions in this interview, you may choose to say "skip," "pass," or "I prefer not to answer." There will be no repercussions for choosing not to answer any of my questions today.

Do you have any questions for me? *Interviewer should answer any questions. If there are no questions say,* **"Alright, let's begin."**

NOTE: If a teacher responds with only a "yes," "no," "maybe," or a brief response, ask a follow up question. Follow-ups examples include, "could you explain a little further," "could you expand on that," "why/why not?"

My first set of questions will ask you about your general impressions of the Be Good People program and its usability from the beginning of the year until now.

1. What do you like the most about the Be Good People program?
2. What do you like the least about the Be Good People program?
3. How would you change the Be Good People program to better suit **your** needs, if at all?
4. How would you change the Be Good People program to better suit **your students'** needs, if at all?
5. How do you see your students using Be Good People **skills at school**, if at all?
6. Do you think your students use Be Good People skills **outside of school?** Why or why not?
7. How **relevant** are the Be Good People skills to your students' lives, if at all?
8. How **important** are the Be Good People skills to your students' lives, if at all?

My second set of questions will ask you about your perceptions surrounding the cultural relevance of the Be Good People program from the beginning of the year until now. I will also give you a gentle reminder now that you do not have to answer any questions you do not want to.

To make sure we have the same understanding, I will be giving us a common definition of culture. Culture is passed down from one generation to the next. Things that can get passed down that encompass culture can include values, norms, goals, items, language, and religion. Culture can affect our behavior and our perceptions of others' behaviors.

9. Given this definition of culture, with a particular focus on the aspects that most impact your daily life, how would you describe your culture?
10. How would you say the description of your culture is similar to or differs from the culture of your students?

11. How would you say your culture is reflected in the Be Good People program, if at all?
12. How would you say **all of your** students' cultures are reflected in the Be Good People program, if at all?
13. In what ways is the Be Good People program made for individuals like your students, if at all?
14. In what ways would you change the program to better suit your **students' cultures,** if at all?
15. *Culturally Relevant Pedagogy* is a model that supports student academics, learning of their own and others' cultures, developing an understanding of social, political, and economic contradictions and how to take action against oppression. **With this shared understanding of Culturally Relevant Pedagogy, have you received any training in Culturally Relevant Pedagogy in SEL or otherwise?**
    a. If so, what did it look like?
16. How do you make changes to the BGP curriculum to align with culturally relevant pedagogy, if at all?

Note: The definition of culture used comes from Fiske (2002) and (Spencer-Oatley, 2012). The definition of Culturally Relevant Pedagogy comes from Ladson-Billings (1995) and (Freire, 1970, p. 35).

Lastly, I have a couple of questions about the training you received for Be Good People.

17. Can you describe what your training in Be Good People was like?
18. How satisfied were you with the training in terms of preparing you to teach the curriculum and why?
19. How would you change the Be Good People training, if at all?

SAY: Any last things you want to note about the Be Good People program?

Thank you, again, for your time today! If you have any follow up questions, please feel free to reach out to me.

Appendix J

Example of How the Representation Tally was Conducted

Lesson Name	Slide #: Character Name/Role - gender, race
Following Instructions	7: Jacob - boy ambiguous 8: Kayla - girl ambiguous 9: Jose - boy ambiguous 9: Jose's teacher – girl, ambiguous 10: Mia - girl ambiguous 11: john - boy ambiguous 11: john's teacher – girl ambiguous Three teachers: unknown, ambiguous

For this example lesson, there were a total of four girls who were racially ambiguous, three boys who were racially ambiguous, and three individuals with unknown genders who were racially ambiguous. In the table below, there are the characteristics for which raters were to tally.

Culturally Responsive Curriculum Scorecard – Representation Portion: Diversity of Characters Tally

	Girl/Woman	Boy/Man	Non Binary	Unknown	Total
Middle Eastern					
Asian/ Pacific Islander					
Black/ African					
Latinx					
Native American					
White					
Racially Ambiguous					
Multiracial					
Peoples with Disabilities					
LGBTQIA+					
Animals					
Total Characters Depicted:					